

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 14:17:53 ; Search time 48 Seconds

(without alignments)
188.488 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304

Sequence: 1 MSTKQFNLDIVSVSKDSGA.....ALMGNMKATCHCHSHSK 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
A.Geneseq.19jun03.*
1: /SIDSI/gcgdata/geneq/geneq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneq/geneq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneq/geneq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneq/geneq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneq/geneq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneq/geneq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneq/geneq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneq/geneq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneq/geneq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneq/geneq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneq/geneq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneq/geneq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneq/geneq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneq/geneq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneq/geneq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneq/geneq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneq/geneq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneq/geneq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneq/geneq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneq/geneq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneq/geneq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneq/geneq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneq/geneq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneq/geneq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	298	98.0	57	11 AAR05238	Nisin precursor pr
2	298	98.0	57	17 AAR95267	Pre-nisin A. Lact
3	298	98.0	57	20 AAY06655	Nisin A of lactoba
4	298	98.0	57	20 AAY03208	Amino acid sequenc
5	291	95.7	57	13 AAR28298	Sequence encoded b
6	291	95.7	57	14 AAR41280	Bacteriocin. Lact
7	291	95.7	57	14 AAR33850	Bacteriocin IL-2 p
8	291	95.7	57	14 AAR39112	Lactococcal bacter
9	291	95.7	57	20 AAY06670	Nisin Z of lactoba

10	185	60.9	34	15 AAR62635	Putative intermedi
11	185	60.9	34	19 AAM66443	Cationic peptide n
12	185	60.9	34	21 AAY91742	Cationic peptide N
13	185	60.9	34	23 AAU90978	Transplant media a
14	185	60.9	34	24 ABUS9619	Cationic cancer -t
15	181	59.5	41	18 AAY31659	Subtilin-nisin chi
16	178	58.6	34	14 AAR43070	Lactococcus lactis
17	176	57.9	34	20 AAM84352	Peptide OSP-91241
18	157	51.6	56	11 AAR05237	Subtilin precursor
19	157	51.6	56	11 AAR66445	Cationic peptide s
20	157	51.6	56	20 AAY03209	Amino acid sequenc
21	157	51.6	56	21 AAY91744	Cationic peptide S
22	157	51.6	56	21 ABUS9621	Cationic cancer -t
23	128	42.1	39	18 AAY31658	Nisin A/SSA. Lact
24	105	34.5	34	14 AAR50953	Epidermin. Staphy
25	104	34.2	34	14 AAR37314	Nisin A. Lactococ
26	104	34.2	34	14 AAR43272	Sequence of the ba
27	104	34.2	34	15 AAR56162	Nisin A. Lactococ
28	104	34.2	34	17 AAR95263	Sequence of nisin
29	97	31.9	34	13 AAR28299	Nisin Z. Lactococ
30	97	31.9	34	17 AAR95245	Lactococcus lactis
31	97	31.9	34	23 AAU75775	Lactococcus lactis
32	97	31.9	34	23 AAU75777	Lactococcus lactis
33	96	31.6	34	23 AAR50954	Nisin A/H270. Lac
34	96	31.6	34	23 AAU75776	Lactococcus lactis
35	94.5	31.1	51	24 ABG72550	Streptococcus muta
36	94.5	31.1	52	10 AAR98498	Sequence of pre-pe
37	94.5	31.1	52	14 AAR37315	Epia protein. Sta
38	94.5	31.1	52	20 AAY43430	S. epidermis readi
39	94.5	31.1	52	20 AAY03210	Amino acid sequenc
40	93	30.6	34	21 AAB11032	Lactococcus lactis
41	92	30.3	34	14 AAR39311	Lactococcal bacter
42	82	27.0	34	14 AAR41281	Bacteriocin (Gener
43	82	27.0	34	14 AAR30171	Bacteriocin IL-2.
44	82	27.0	47	16 AAR85080	Hyacin M51 (lanthi
45	81.5	26.8	35	18 AAY31660	Mutant lantibiotic

ALIGNMENTS

RESULT 1	AA05238	standard; protein; 57 AA.
ID	AA05238	
AC	AA05238;	
XX		
DT	04-AUG-1990 (first entry)	
DE	Nisin precursor protein and leader peptide encoded by DNA derived from	
DE	Streptococcus lactis ATCC 11454.	
XX		
KW	Nisin precursor peptide; nisin leader peptide;	
KW	post-translational modification; Streptococcus lactis ATCC 11454;	
XX		
OS	Streptococcus lactis ATCC 11454.	
XX		
FM	Key	Location/Qualifiers
FT	Peptide	1..23
FT		/note="leader fragment responsible for inducing post-
FT		translational modification"
FT	Protein	24..57
FT		/note="nisin precursor peptide"
PN	WO9000558-A.	
XX		
PD	25-JAN-1990.	
XX		
PF	30-JUN-1989;	89WO-US02820.
XX		
PR	05-JUL-1988;	88US-0214959.
XX		
PA	(UWVA-) UNIV OF MARYLAND.	

```

XX Hansen NJ,
XX
XX WPI: 1990-051685/07.
XX N-PSDB; AAO33354.
XX
XX Leader peptide sequence -
XX including post-translational modification of polypeptide(s)
XX
XX Disclosure; Fig 3; 19pp; English.
XX
XX The leader peptide assists in inducing post-translational modification in
XX a protein precursor when attached to the precursor as a leader. The
XX precursor polypeptide contains Ser, Thr and Cys which undergo
XX modification after translation to arrive at the mature protein,
XX having unusual amino acids. Tag a is claimed in the patent.
XX
XX Sequence 57 AA;
XX
XX Query Match 98.0%; Score 298; DB 11; Length 57;
XX Best Local Similarity 98.2%; Pred. No. 9.9e-29;
XX Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDLVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 MSTKDFNLDLVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 2
AAR95267
ID AAR95267 standard; Protein; 57 AA.
XX
XX AAR95267;
XX
XX 01-AUG-1996 (first entry)
XX
XX -Pre-nisin A.
XX
XX Nisin A; nisa gene; antimicrobial; preservative; antibiotic;
XX lantibiotic; protein engineering.
XX
XX Lactococcus lactis strain NIZO R5.
XX
XX OS
XX FH Key Location/Qualifiers
XX FT Peptide 1..23
XX FT /label= Sig_peptide
XX
XX WO9616180-A1.
XX
XX PD 30-MAY-1996.
XX
XX PF 20-NOV-1995; 95WO-GB02699.
XX
XX PR 19-NOV-1994; 94GB-0023404.
XX
XX PA (BIOT-) BIOTECHNOLOGY & BIOLOGICAL SCI RES COUNC.
XX
XX Dadd HM, Gasson MJ;
XX
XX WPI: 1996-268616/27.
XX
XX N-PSDB; AAT29660.
XX
XX Making cell which expresses nisa but does not contain natural nisa
XX gene - by providing cell with variant nisa gene, and genes for nisa
XX modification, secretion and immunity
XX
XX Disclosure; Fig 7; 69pp; English.
XX
XX The gene cluster nisaBTPK (see AAT29660 and AAT29661) of Lactococcus
XX lactis includes the nisa gene coding for pre-nisin A (AAR95267, see
XX also AAR95263) and the genes for nisa modification, secretion and
XX immunity. nisa (AAR95268) and nisc (AAR95270) are believed to be
XX involved in reactions that modify pre-nisin; nist (AAR95269) is

```

```

CC similar to a transport Appase and is involved in translocation of
CC nisin out of the cell; nist (AAR95271) is involved in immunity to
CC nisin. Replacement of the natural, chromosomal copy of the nisa
CC gene with a variant nisa gene allows prodn. of high levels of nisin
CC A variants in Lactococcus lactis hosts.
XX
XX Sequence 57 AA;
XX
XX Query Match 98.0%; Score 298; DB 17; Length 57;
XX Best Local Similarity 98.2%; Pred. No. 9.9e-29;
XX Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDLVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 MSTKDFNLDLVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 3
AA06665
ID AA06665 standard; Protein; 57 AA.
XX
XX AA06665;
XX
XX 09-NOV-1999 (first entry)
XX
XX Nisin A of Lactobacillus lactis.
XX
XX Nisin A; nisa gene; antimicrobial; animal performance; growth rate;
XX feed conversion; probiotic; bacteriocin.
XX
XX Lactobacillus lactis.
XX
XX OS
XX PN WO9941978-A1.
XX
XX PD 26-AUG-1999.
XX
XX PF 12-FEB-1999; 99WO-IB00250.
XX
XX PR 18-FEB-1998; 98GB-0003424.
XX
XX PA (PF12) PFIZER INC.
XX
XX PA (PF12) PFIZER LTD.
XX
XX Flanagan AJ, Haxell MA, Rolph TP;
XX
XX WPI: 1999-527402/44.
XX
XX N-PSDB; AAX87792.
XX
XX Novel performance enhancing method resulting in improved growth
XX rates and feed conversion efficiencies in animals
XX
XX Disclosure; Page 61; 79pp; English.
XX
XX The present sequence represents the nisa A precursor peptide of
XX Lactobacillus lactis NIZO R5. The sequence is deduced from the
XX nisa A gene (see AAX87792). Nisin A is lantibiotic-containing
XX bacteriocin. The invention relates to methods of enhancing
XX performance in an animal by administering a bacterium capable of
XX expressing a performance enhancing polypeptide such as nisa A or
XX nisin Z. The enhanced performance results in improved growth rates
XX and feed conversion efficiencies. The bacterium, which may be
XX genetically modified to express the performance enhancing
XX polypeptide, is administered to the gastrointestinal tract,
XX especially to an embryo or neonatal animal. If the polypeptide is
XX nisin, it may also inhibit ruminal methane, decrease acetate to
XX propionate ratios and prevent amino acid deamination.
XX Administration of an appropriate bacterium which may act as a
XX probiotic may also help control enteric pathogens in poultry. The
XX polypeptide is produced continuously in the gut, maintaining a
XX constant level. The protein is eventually degraded, leaving no
XX residues in the meat.
XX
XX Sequence 57 AA;

```

Query Match 98.0%; Score 298; DB 20; Length 57;
 Best Local Similarity 98.2%; Pred. No. 9.9e-29;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
 DB 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 4

ID AAY03208 standard; Protein; 57 AA.

AC AAY03208;

DT 03-AUG-1999 (first entry)

DE Amino acid sequence of nisin A.

KM Sublancin 168; antimicrobial; food preservative; pro-sublancin 168;
 KW Gram-positive bacteria; pre-sublancin 168; nisin A.

OS unknown.

PN WO9903352-A1.

PD 28-JAN-1999.

PF 17-JUL-1998; 98WO-US14547.

PR 18-JUL-1997; 97US-0053035.

PA (UTMA-) UNIV MARYLAND BALTIMORE.

PI Hansen JN;

DR WPI; 1999-131752/11.

PT New antimicrobial peptide, sublancin 168, from *Bacillus subtilis* -
 used for, e.g. treatment of infections caused by Gram negative
 bacteria and as food preservative

PS Disclosure; Page 53; 71pp; English.

CC This is the amino acid sequence of nisin A used in the method of
 the invention involving the use of prosublancin 168. The peptide
 designated sublancin 168, is an antimicrobial useful for treating
 CC infections and preserving food against spoilage bacteria.
 CC particularly Gram-positive bacteria. Pro-sublancin 168 and
 CC pre-sublancin 168, are the precursors of sublancin 168. Sublancin
 CC 168 is very stable at low pH and can be autoclaved without damage.
 CC It does not decompose after 2 years in aqueous solution of about
 CC neutral pH.

XX Sequence 57 AA;

Query Match 98.0%; Score 298; DB 20; Length 57;
 Best Local Similarity 98.2%; Pred. No. 9.9e-29;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
 DB 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 5

ID AAR28298 standard; Protein; 57 AA.

AC AAR28298;

DT 25-MAR-2003 (updated)

DT 02-APR-1993 (first entry)

DE Séquencé encoded by nisz gene isolated from *L. Lactis* NIZO 22186.

KM Lantibiotic; nisin Z; nisin A; analogue; food preservative.

OS Lactococcus lactis.

FT Key Location/Qualifiers

FT Peptide 1..23
 /label= leader

PN WO9218633-A1.

PD 29-OCT-1992.

PF 09-APR-1992; 92WO-NL00068.

PR 11-APR-1991; 91NL-0000634.

PA (NIZO-) NIZO STICHTING NEDERLANDS INST ZUIVELOND.

PI De Vos WM, Kuipers OP, Siezen RJ;

DR WPI; 1992-382116/46.

DR N-PSDB; AAQ28299.

PT New lantibiotic cpds. related to nisin A - and *Lactococcus*
 PT strains which produce them, useful as preservatives for foods and
 PT animal feeds

PS Example; Fig 3; 42pp; English.

CC Total DNA was isolated from the *L. lactis* strains NIZO 22186 and
 CC NIZO R5. With the aid of the nisa gene of strain NIZO R5 as a probe,
 CC a 4.5 kb HindIII fragment was identified in the total DNA strain
 CC NIZO 22186, which was then cloned in M13 mp18. The DNA sequence of
 CC the gene for nisin Z production (nisZ) was determined by making use
 CC of oligos complementary to the 5' and 3' flanking sequences of the
 CC nisa gene. The nucleotide sequence of the nisz gene is found to be
 CC identical to that of the nisa gene with the exception of a C to A
 CC transversion in posn. 148 which results in the replacement of AA
 CC His27 by Asn27. The above indicate that the structure of nisin Z is
 CC as shown in AAR28299.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 57 AA;

Query Match 95.7%; Score 291; DB 13; Length 57;
 Best Local Similarity 96.5%; Pred. No. 7e-28;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
 DB 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 6

ID AAR41280 standard; Protein; 57 AA.

AC AAR41280;

DT 25-MAR-2003 (updated)

DT 10-MAR-1994 (first entry)

DE Bacteriocin.

KM Bacteriocin; LL-2; gram positive bacteria; antimicrobial; food;
 KW inhibit.

OS *Lactococcus lactis*.

```

FH Key Location/Qualifiers
FT Peptide 1..23
FT /label= sig_peptide
FT 24..57
FT Protein /label= mat_protein
FT /note= "Claim 3"
FT Modified-site 26
FT /note= "The CH2 in the side chain joins
FT the S in residue 30"
FT Modified-site 30
FT /note= "The S joins the side chain of residue 26"
FT Modified-site 31
FT /note= "The CH2 in the side chain joins
FT the S in residue 34"
FT Modified-site 34
FT /note= "The S joins the side chain of residue 31"
FT Modified-site 36
FT /note= "The CH2 in the side chain joins
FT the S in residue 42"
FT Modified-site 42
FT /note= "The S joins the side chain of residue 36"
FT Modified-site 46
FT /note= "The CH2 in the side chain joins
FT the S in residue 51"
FT Modified-site 51
FT /note= "The S joins the side chain of residue 46"
FT Modified-site 49
FT /note= "The CH2 in the side chain joins
FT the S in residue 52"
FT Modified-site 52
FT /note= "The S joins the side chain of residue 49"
PN US5232849-A.
PD 03-AUG-1993.
XX -14-MAY-1992; 92US-0882079.
XX PF
XX -01-JUL-1991; 91US-0721774.
XX PR
XX 14-MAY-1992; 92US-0882079.
XX PA
XX (UNIL ) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
XX Henderson JT, Marug JD, Van Wassenar PD, Vedamuthu ER;
XX MPI; 1993-287077/36.
XX DR N-PSDB; AAQ49150.
XX
XX Bacteriocin from Lactococcus lactis subspecies lactis - useful as
XX inhibitory against Gram-positive bacteria
XX
XX Disclosure; Page 13-14 (col 15,16,17,18); 14pp; English.
XX
XX The sequence (AAQ49150) was amplified using primers (AAQ49151-52). The
XX encoded protein inhibits selected gram positive bacteria and this
XX property is enhanced if further purified by HPLC. The materials
XX being treated to provide inhibition are preferably foods, although
XX other materials may be treated.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 57 AA;

```

```

Query Match 95.7%; Score 291; DB 14; Length 57;
Best Local Similarity 96.5%; Pred. No. 7e-28;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MSTKDFNLDIVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHGSIHVSK 57
DB 1 MSTKDFNLDIVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHGSIHVSK 57

```

RESULT 7

```

AAR33850
ID AAR33850 standard; Protein; 57 AA.
XX
XX AAR33850;
AC
XX 25-MAR-2003 (updated)
DT 12-MAY-1993 (first entry)
XX
DE Bacteriocin LL-2 precursor.
XX
XX Food treatment.
XX
XX Lactococcus lactis sub-species lactis NRRL B-18809.
XX
XX
XX Key Location/Qualifiers
FT Peptide 1..23
FT /note= "signal peptide"
FT 24..57
FT Peptide /*note= "mature peptide"
FT
XX
XX US5173297-A.
XX PN
XX 22-DEC-1992.
XX PD
XX 01-JUL-1991; 91US-0721774.
XX PF
XX 01-JUL-1991; 91US-0721774.
XX PR
XX 01-JUL-1991; 91US-0721774.
XX PA
XX (UNIL ) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
XX Vedamuthu ER, Henderson JT, Marug JD, Vanwassenar PD;
XX MPI; 1993-017533/02.
XX DR N-PSDB; AAQ34782.
XX
XX Inhibition of Gram-positive bacteria - using bacteriocin derived
XX from Lactococcus lactis sub-species lactis NRRL B-18809
XX
XX Disclosure; Page 13; 14pp; English.
XX
XX The sequence is that of bacteriocin LL-2 precursor which can be
XX used in a method for the inhibition of Gram-positive bacteria.
XX LL-2 is especially useful for treatment of food, although other
XX non-food materials may also be treated.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 57 AA;

```

```

Query Match 95.7%; Score 291; DB 14; Length 57;
Best Local Similarity 96.5%; Pred. No. 7e-28;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MSTKDFNLDIVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHGSIHVSK 57
DB 1 MSTKDFNLDIVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHGSIHVSK 57

```

```

RESULT 8
AAR39312
ID AAR39312 standard; Protein; 57 AA.
XX
XX AAR39312;
AC
XX 25-MAR-2003 (updated)
DT 21-JAN-1994 (first entry)
XX
XX Lactococcal bacteriocin polypeptide precursor.
XX Bacteriocin; inhibition; polypeptide; Lactococcus lactis.
XX Lactococcus lactis (subspecies lactis).
XX

```

```

XX Lactococcus lactis (subspecies lactis).
XX

```


FH Key Location/Qualifiers
 FT Protein 24.57
 XX /label= Bacteriocin.
 XX
 XX US5231165-A.
 XX
 XX 27-JUL-1993.
 XX
 XX 14-MAY-1992; 92US-0882715.
 XX
 XX 01-JUL-1991; 91US-0721774.
 XX
 XX 14-MAY-1992; 92US-0882715.
 XX
 XX (UNTIL) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
 XX
 XX Henderson JT, Marugg JD, Vanwassenaar PD, Vedamuthu ER;
 XX
 XX WPI; 1993-249768/31.
 XX
 XX N-PSDB; AAQ46818.
 XX
 XX Isolated and purified polypeptide from *Lactococcus lactis* sub
 PT species *lactis* - has inhibitory activity against gram-positive
 PT bacteria for e.g. food etc.
 XX
 XX Claim 1; Column 17-18; 13pp; English.
 XX
 XX The isolated bacteriocin obtained from the polypeptide precursor has
 CC an inhibitory activity against selected gram positive bacteria.
 CC The amount of bacteriocin required to provide inhibition is 15-100
 CC arbitrary units per gram of material. The materials being treated
 CC with the bacteriocin to provide inhibition are especially foodstuffs.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 XX Sequence 57 AA;
 SQ
 Query Match 95.7%; Score 291; DB 14; Length 57;
 Best Local Similarity 96.5%; Pred. No. 7e-28;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSTKDFNLDLVSVSKDGSAPRITSTSLCTPGCKTGALMGCMKMTATCNCISIVSK 57
 DB 1 MSTKDFNLDLVSVSKDGSAPRITSTSLCTPGCKTGALMGCMKMTATCNCISIVSK 57
 RESULT 9
 AAY06670
 ID AAY06670 standard; Protein; 57 AA.
 XX
 XX AAY06670;
 AC
 XX
 XX 09-NOV-1999 (first entry)
 DT
 XX
 XX Nisin Z of *Lactobacillus lactis*.
 DE
 XX
 XX Nisin Z; nis Z; lantibiotic; animal performance; growth rate;
 KM feed conversion; bacteriocin; probiotic.
 KW
 XX
 XX *Lactobacillus lactis*.
 OS
 XX
 XX WO9941978-A1.
 PN
 XX
 XX 26-AUG-1999.
 PD
 XX
 XX 12-FEB-1999; 99WO-1B00250.
 PF
 XX
 XX 18-FEB-1998; 98GB-0003424.
 PR
 XX
 XX (PFIZ) PFIZER INC.
 PA (PFIZ) PFIZER LTD.
 PA
 XX
 XX Flanagan AJ, Haxell MA, Rolph TP;
 PI
 XX

DR WPI; 1999-527402/44.
 DR N-PSDB; AAX87793.
 XX
 XX Novel performance enhancing method resulting in improved growth
 PT rates and feed conversion efficiencies in animals
 PT
 XX
 XX Disclosure; Page 71; 79pp; English.
 XX
 XX This sequence represents the nisin Z peptide of *Lactobacillus*
 CC *lactis* strain 22186. Nisin Z is a natural analogue of nisin A (see
 CC AAY06665), a lantibiotic-containing bacteriocin. The invention
 CC relates to methods of enhancing performance in an animal by
 CC administering a bacterium capable of expressing a performance
 CC enhancing polypeptide such as nisin A or nisin Z. The enhanced
 CC performance results in improved growth rates and feed conversion
 CC efficiencies. The bacterium, which may be genetically modified to
 CC express the performance enhancing polypeptide, is administered to
 CC the gastrointestinal tract, especially to an embryo or neonatal
 CC animal. If the polypeptide is nisin, it may also inhibit ruminal
 CC methane, decrease acetate to propionate ratios and prevent amino
 CC acid deamination. Administration of an appropriate bacterium which
 CC may act as a probiotic may also help control enteric pathogens in
 CC poultry. The polypeptide is produced continuously in the gut,
 CC maintaining a constant level. The protein is eventually degraded,
 CC leaving no residues in the meat.
 XX
 XX Sequence 57 AA;
 SQ
 Query Match 95.7%; Score 291; DB 20; Length 57;
 Best Local Similarity 96.5%; Pred. No. 7e-28;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSTKDFNLDLVSVSKDGSAPRITSTSLCTPGCKTGALMGCMKMTATCNCISIVSK 57
 DB 1 MSTKDFNLDLVSVSKDGSAPRITSTSLCTPGCKTGALMGCMKMTATCNCISIVSK 57
 RESULT 10
 AAR62635
 ID AAR62635 standard; peptide; 34 AA.
 XX
 XX AAR62635;
 AC
 XX
 XX 15-JUN-1995 (first entry)
 DT
 XX
 XX Putative intermediate for lantionine-contg. peptide.
 DE
 XX
 XX lantionine; methylantionine; lantibiotic; antiviral;
 KW immunosuppressant; antimicrobial; enzyme inhibitor.
 XX
 XX Synthetic.
 OS
 XX
 XX JP06253885-A.
 PN
 XX
 XX 13-SEP-1994.
 PD
 XX
 XX 09-MAR-1993; 93JP-0048385.
 PF
 XX
 XX 09-MAR-1993; 93JP-0048385.
 PR
 XX
 XX (AJIN) AJINOMOTO KK.
 PA
 XX
 XX WPI; 1994-329026/41.
 DR
 XX
 XX Prepn. of lantionine contg. peptide(s) - useful as antimicrobial,
 PT antiviral drugs, immunosuppressants and enzyme inhibitors
 PT
 XX
 XX Example 2; Page 7; 8pp; Japanese.
 PS
 XX
 XX This is one of 5 peptides (AAR62635-R63639) containing Cys and Ser or
 CC Thr residues which were synthesised and tested for their usefulness
 CC as intermediates for the preparation of peptides which include
 CC lantionine. Peptides 3 and 4 (AAR62637-8) produced lantionine, while

CC both lanthionine and methylanthionine could be produced from peptide
 CC 2 (AAR62636).

XX
 SQ Sequence 34 AA;

Query Match 60.9%; Score 185; DB 15; Length 34;
 Best Local Similarity 97.1%; Pred. No. 2.8e-15;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 24 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
 DB 1 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 34

RESULT 11
 ID AAW66443
 AA66443 standard; peptide; 34 AA.

XX
 AC AAW66443;

XX
 DT 12-JAN-1999 (first entry)

XX
 DE Cationic peptide nisin.

XX
 KW Indolicidin analogue; resistance; cationic peptide; antibiotic;

XX
 KM bacterial infection; tolerance; antibacterial; microorganism;

XX
 OS bacteria; fungus; parasite; virus.

XX
 PN Lactococcus lactis.

XX
 PD WO9840401-A2.

XX
 PD 17-SEP-1998.

XX
 PF 10-MAR-1998; 98WO-CA00190.

XX
 PR -25-FEB-1998; 98US-0030619.

XX
 PR 10-MAR-1997; 97US-0040649.

XX
 PR 20-AUG-1997; 97US-0915314.

XX
 PR 26-SEP-1997; 97US-0060099.

XX
 PA (MICR-) MICROLOGIX BIOTECH INC.

XX
 PI Fraser JR, McNicol PJ, West MHP;

XX
 DR WPI; 1998-520800/44.

XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing

XX
 PT activity of antibiotic or overcoming tolerance, acquired resistance

XX
 PT or inherent resistance of microorganisms

XX
 PS Disclosure; Page 10; 105pp; English.

XX
 CC AAW66393 to AAW66469 represent native cationic peptides from the

XX
 CC present invention. The present invention describes compositions and

XX
 CC methods for treating infection, especially bacterial infections. The

XX
 CC compositions and methods use cationic peptides in combination with an

XX
 CC antibiotic agent which are then administered to a patient to enhance the

XX
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)

XX
 CC acquired resistance; and (c) inherent resistance. The combinations of

XX
 CC antibiotics and cationic peptides can provide synergistic activity

XX
 CC against a microorganism that is tolerant, inherently resistant, or has

XX
 CC acquired resistance to an antibiotic agent. They can be used for killing

XX
 CC e.g. bacteria, fungi, parasites and viruses.

XX
 SQ Sequence 34 AA;

XX
 Query Match 60.9%; Score 185; DB 19; Length 34;

XX
 Best Local Similarity 97.1%; Pred. No. 2.8e-15;

XX
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX
 OY 24 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

XX
 DB 1 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 34

XX
 RESULT 12

XX
 ID AAY91742 standard; Peptide; 34 AA.

XX
 AC AAY91742;

XX
 DT 06-JUN-2000 (first entry)

XX
 DE Cationic peptide Nisin amino acid sequence.

XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;

XX
 KM leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;

XX
 KM breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;

XX
 KM multidrug resistance.

XX
 OS Unidentified.

XX
 PN WO9965506-A2.

XX
 PD 23-DEC-1999.

XX
 PF 14-JUN-1999; 99WO-CA00552.

XX
 PR 12-JUN-1998; 98US-0096541.

XX
 PA (MICR-) MICROLOGIX BIOTECH INC.

XX
 PI Friedland HD, Krieger TJ, Taylor R, Erffle D, Fraser JR, West MHP;

XX
 DR WPI; 2000-223549/19.

XX
 PT Novel pharmaceutical composition containing optionally activated

XX
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours

XX
 PS Disclosure; Page 11; 94pp; English.

XX
 CC This sequence represents a cationic peptide amino acid sequence, which

XX
 CC can be used in the pharmaceutical composition of the invention. The

XX
 CC invention relates to a pharmaceutical composition containing at least one

XX
 CC activated polyoxalkylene (APO)-modified cationic peptide. The

XX
 CC modification of peptides with APO increases their activity against tumour

XX
 CC cells, including those with a multidrug resistant phenotype. The

XX
 CC pharmaceutical composition can be used to treat tumours, specifically

XX
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,

XX
 CC cervix, uterus, skin, prostate, liver and colon.

XX
 SQ Sequence 34 AA;

XX
 Query Match 60.9%; Score 185; DB 21; Length 34;

XX
 Best Local Similarity 97.1%; Pred. No. 2.8e-15;

XX
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX
 OY 24 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

XX
 DB 1 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 34

XX
 RESULT 13

XX
 ID AAU90978 standard; Peptide; 34 AA.

XX
 AC AAU90978;

XX
 DT 05-JUN-2002 (first entry)

XX
 DE Transplant media associated antimicrobial peptide #14.

XX
 KM Transplant; antimicrobial peptide; pore forming agent;

XX
 KM cell surface receptor binding compound; kidney transplant;

KW cardioplegia; organ transplant; transplant rejection.
 OS Lactococcus lactis.
 XX WO200209738-A1.
 XX
 XX 07-FEB-2002.
 PD
 XX 27-JUL-2001; 2001WO-US23785.
 PF
 XX 28-JUL-2000; 2000US-221632P.
 PR 17-NOV-2000; 2000US-249602P.
 PR 15-MAY-2001; 2001US-290932P.
 XX
 XX (MURP/) MURPHY C J.
 PA
 XX Murphy CJ, Reid TW, Meanlty JF;
 PI WPI; 2002-268995/31.
 DR
 XX Media comprising antimicrobial polypeptides or pore forming agents
 PT and/or cell surface receptor binding compounds useful for the storage
 PT and preservation of organs prior to transplant -
 PS Disclosure; Page 25; 78pp; English.
 XX
 XX The invention describes new transplant compositions comprising
 CC antimicrobial polypeptides or pore forming agents and/or cell surface
 CC receptor binding compounds. The media is capable of extending the
 CC preservation period past 72 hours and can provide organs with increased
 CC functionality upon transplant. animals receiving kidneys stored in the
 CC media of the present invention for either three or four days had serum
 CC creatinine levels of less than half of those observed in control animals
 CC receiving kidneys stored in UW solution (defined in the specification)
 CC alone. Lower serum creatinine levels are indicative of healthier kidneys
 CC and a more preferable prognosis for the transplant patient. The media of
 CC the invention are useful for decreasing the incidence and/or severity of
 CC delayed graft function in patients receiving transplanted kidneys stored
 CC and/or treated in the media. The media may also be used in procedures
 CC such as cardioplegia. It is contemplated that transplant of healthier
 CC organs leads to a decrease in chronic rejection. This sequence represents
 CC an antimicrobial peptide studied in the development of the transplant
 CC media.
 CC
 SQ Sequence 34 AA;
 Query Match 60.9%; Score 185; DB 23; Length 34;
 Best Local Similarity 97.1%; Pred. No. 2.8e-15;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 24 ITSTSLCTPGCKTGTALMGCMKMTATCHCSIHVSK 57
 Db 1 ITSTSLCTPGCKTGTALMGCMKMTATCHCSIHVSK 34
 RESULT 14
 AAY31659
 ID AAY31659 standard; Peptide; 34 AA.
 AC
 XX ABUS9619;
 XX
 XX 22-APR-2003 (first entry)
 DE Cationic cancer -targeting peptide #51.
 XX
 XX Targeting ligand; bioactive agent; polymer matrix; cancer; cytostatic;
 KW cathepsin-D substrate; peptides; neuroreceptor; adrena receptor;
 KW fibronectin; vitronectin; integrin; RGD motif; angiogenic endothelium;
 KW tumour; cationic cancer-targeting peptide.
 XX
 XX Synthetic.
 OS
 XX US2002041898-A1.
 PN

XX
 PD 11-APR-2002.
 XX
 XX 25-JUL-2001; 2001US-0912609.
 PF
 XX 05-JAN-2000; 2000US-0478124.
 PR 31-OCT-2000; 2000US-0703474.
 XX
 XX (UNGE/) UNGER E C.
 PA (MATS/) MATSUNAGA T O.
 PA (RAMA/) RAMASWAMI V.
 PA (ROMA/) ROMANOWSKI M J.
 XX
 PI Unger EC, Matsunaga TO, Ramaswami V, Romanowski MJ;
 DR WPI; 2003-208921/20.
 XX
 XX Targeted delivery system comprising a bioactive agent homogeneously
 PT dispersed in a targeted matrix is especially useful in cancer therapy
 PT
 PS Disclosure; Page14; 46pp; English.
 XX
 XX The invention relates to a composition comprising a bioactive agent
 CC homogeneously dispersed in a targeted matrix (polymer and targeting
 CC ligand). Also included are a targeted matrix for use as a delivery
 CC vehicle comprising a polymer associated with a targeting ligand,
 CC enhancing the bioavailability of an agent comprising administration of the
 CC of the composition and treating cancer comprising administration of a drug,
 CC novel composition. The method is useful for targeted delivery of a drug,
 CC especially in cancer therapy. The targeting ligand may be a peptide.
 CC Examples of targeting peptides are disclosed including cathepsin-D
 CC substrate peptides, peptides targeting receptors in the brain and
 CC kidney, peptides recognising fibronectin- and vitronectin-binding
 CC integrins, peptides targeting the RGD (Arg-Gly-Asp)-motif in, e.g.,
 CC antibodies, peptides targeting the angiogenic endothelium of solid
 CC tumours, tissue specific peptides (e.g. of lung, skin, pancreas,
 CC intestine, uterus, adrenal gland and retina), and cationic cancer-
 CC targeting peptides. The present sequence is a peptide targeting
 CC ligand disclosed in the invention.
 CC
 SQ Sequence 34 AA;
 Query Match 60.9%; Score 185; DB 24; Length 34;
 Best Local Similarity 97.1%; Pred. No. 2.8e-15;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 24 ITSTSLCTPGCKTGTALMGCMKMTATCHCSIHVSK 57
 Db 1 ITSTSLCTPGCKTGTALMGCMKMTATCHCSIHVSK 34
 RESULT 15
 AAY31659
 ID AAY31659 standard; Protein; 41 AA.
 AC
 XX AAY31659;
 XX
 XX 09-NOV-1999 (first entry)
 DE Subtilin-nisin chimera.
 XX
 XX Nisin; subtilin; lantibiotic; chimera; mutant; bacteriocide;
 KW preservative.
 XX
 XX Chimeric - Lactococcus lactis.
 OS Chimeric - Bacillus subtilis.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..7
 FT /note= "signal peptide"
 FT 8..41
 FT Protein /note= "mature protein"

```

FT Region 8..18
FT /note= "nisin (1-11)"
FT Region 19..41
FT /note= "subtilin(12-32)"
XX
XX PN WO9711713-A1.
XX
XX PD 03-APR-1997.
XX
XX PF 30-SEP-1996; 96WO-US15160.
XX
XX PR 28-SEP-1995; 95US-0535494.
XX
XX PA (UYWA-) UNIV MARYLAND BALTIMORE.
XX
XX PI Hansen JN;
XX
XX DR WPI; 1997-225847/20.
XX DR N-PSDB; AAX87829.
XX
XX PT Lantibiotic mutants and chimera(s) - having enhanced stability and
XX PT activity compared to nisin
XX
XX PS Example; Fig 2; 60pp; English.
XX
XX CC The present sequence represents a chimeric pre-peptide composed of
XX CC a subtilin leader region and a subtilin-nisin fusion comprising
XX CC residues 1-11 of Bacillus subtilis subtilin and residues 12-37 of
XX CC Lactococcus lactis nisin. The subtilin-nisin fusion was not
XX CC processed into a functional lantibiotic when expressed in B.
XX CC subtilis. A heterogeneous mixture of products was produced, with
XX CC none of the products having the expected properties of a correctly
XX CC processed polypeptide. However, the mixture contained a minor
XX CC component with a specific activity that exceeded that of nisin.
XX CC The invention provides lantibiotic mutants and chimeras (see also
XX CC AAY31658) having enhanced activity and stability compared to nisin
XX CC and subtilin. They can be produced by cultivation of transformed
XX CC host cells and used e.g. as food preservatives to treat, kill or
XX CC inhibit the growth of microorganisms and/or their spores.
XX
XX SQ Sequence 41 AA;

```

Query Match 59.5%; Score 181; DB 18; Length 41;
 Best Local Similarity 80.5%; Pred. No. 1.1e-14;
 Matches 33; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

QY 17 DSGASPRITSTLCTPGCKTGALMGCMKMTATCHGSIHVSX 57
DB 1 DSKITPQWKSESPCTPGCKTGALMGCMKMTATCHGSIHVSX 41

```

Search completed: January 12, 2004, 14:24:21
 Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 12:06:16 ; Search time 21 Seconds

(without alignments)
114.844 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304
Sequence: 1 MSTDFNLDIVSVSKDGA.....ALMGCMKATCHGSIHVK 57

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	298	98.0	57	4	US-08-836-687B-20
2	298	98.0	57	4	US-09-462-478A-8
3	295	97.0	57	4	US-08-836-687B-24
4	295	97.0	57	4	US-08-836-687B-26
5	292	96.1	57	4	US-08-836-687B-28
6	291	95.7	57	1	US-08-129-151A-2
7	291	95.7	57	4	US-08-836-687B-40
8	291	95.7	57	4	US-08-836-687B-43
9	291	95.7	57	4	US-08-836-687B-42
10	289	95.1	57	4	US-08-836-687B-46
11	289	95.1	57	4	US-08-836-687B-44
12	288	94.7	57	4	US-08-836-687B-45
13	285	93.8	83	3	US-08-773-731A-2
14	285	93.8	83	3	US-08-773-731A-2
15	282.5	92.9	56	4	US-08-836-687B-41
16	280	92.1	57	4	US-08-836-687B-47
17	185	60.9	34	4	US-09-030-619-206
18	185	60.9	34	6	5231013-4
19	181	59.5	41	2	US-08-535-494-9
20	181	59.5	41	3	US-09-097-635-9
21	178	58.6	34	1	US-07-880-003-1
22	178	58.6	34	2	US-08-030-911-2
23	176	57.9	34	2	US-08-030-911-1
24	157	51.6	56	1	US-07-981-525-2
25	157	51.6	56	1	US-07-981-525-7
26	157	51.6	56	1	US-08-220-033-2
27	157	51.6	56	1	US-08-220-033-7

28	157	51.6	56	2	US-08-465-491-2	Sequence 2, Appli
29	157	51.6	56	2	US-08-465-491-7	Sequence 7, Appli
30	157	51.6	56	2	US-08-986-617-2	Sequence 2, Appli
31	157	51.6	56	2	US-08-986-617-7	Sequence 7, Appli
32	157	51.6	56	4	US-09-030-619-208	Sequence 208, App
33	157	51.6	56	4	US-09-462-478A-9	Sequence 9, Appli
34	154	50.7	56	1	US-07-981-525-9	Sequence 9, Appli
35	154	50.7	56	1	US-08-220-033-9	Sequence 9, Appli
36	154	50.7	56	2	US-08-465-491-9	Sequence 9, Appli
37	154	50.7	56	2	US-08-986-617-9	Sequence 9, Appli
38	128	42.1	39	2	US-08-535-494-5	Sequence 5, Appli
39	128	42.1	39	3	US-09-097-635-5	Sequence 10, Appli
40	113	37.2	23	1	US-08-524-677-10	Sequence 25, Appli
41	113	37.2	23	2	US-08-465-491-25	Sequence 25, Appli
42	113	37.2	23	2	US-08-986-617-25	Sequence 49, Appli
43	106	34.9	34	4	US-08-836-687B-49	Sequence 50, Appli
44	105	34.5	34	4	US-08-836-687B-50	Sequence 51, Appli
45	105	34.5	34	4	US-08-836-687B-51	

ALIGNMENTS

RESULT 1
US-08-836-687B-20
; Sequence 20, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 207477/70
; CURRENT APPLICATION NUMBER: US/08/836, 687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-20

Query Match 98.0%; Score 298; DB 4; Length 57;
Best Local Similarity 98.2%; Pred. No. 1.8e-29;
Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 MSTDFNLDIVSVSKDGAASPRITSTSLCTPGCKTGALMGCMKATCHGSIHVK 57
Db 1 MSTDFNLDIVSVSKDGAASPRITSTSLCTPGCKTGALMGCMKATCHGSIHVK 57

RESULT 2
US-09-462-478A-8
; Sequence 8, Application US/09462478A
; Patent No. 6541607
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF MARYLAND
; TITLE OF INVENTION: SUBSTANTIAL ANTIBIOTIC PRODUCED BY BACILLUS SUBTILIS 168
; FILE REFERENCE: 8172-8072
; CURRENT APPLICATION NUMBER: US/09/462, 478A
; CURRENT FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: US/60/053, 035
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Nisin A
US-09-462-478A-8

Query Match 98.0%; Score 298; DB 4; Length 57;
Best Local Similarity 98.2%; Pred. No. 1,8e-29;
Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 3

US-08-836-687B-24
; Sequence 24, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-24

Query Match 97.0%; Score 295; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 4,1e-29;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 4

US-08-836-687B-26
; Sequence 26, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-26

Query Match 97.0%; Score 295; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 4,1e-29;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVAK 57

RESULT 5

US-08-836-687B-28
; Sequence 28, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN

; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.

US-08-836-687B-28

Query Match 96.1%; Score 292; DB 4; Length 57;
Best Local Similarity 94.7%; Pred. No. 9,6e-29;
Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 MSTKDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVAK 57

RESULT 6

US-08-129-151A-2
; Sequence 2, Application US/08129151A
; Patent No. 5594103
; GENERAL INFORMATION:
; APPLICANT: DE VOS, Willem M.
; APPLICANT: KUIPERS, Oscar P.
; TITLE OF INVENTION: LANTIBIOTICS SIMILAR TO NISIN A, LACTIC
; TITLE OF INVENTION: ACID BACTERIA WHICH PRODUCE SUCH LANTIBIOTICS, METHOD FOR
; TITLE OF INVENTION: CONSTRUCTING SUCH LACTIC ACID BACTERIA AND METHOD FOR
; TITLE OF INVENTION: PRESERVING FOODSTUFFS WITH ETC.
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESS: c/o YOUNG & THOMPSON
; STREET: 745 South 23rd Street, Second Floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/129,151A
; FILING DATE: 07-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 9100634
; FILING DATE: 11-APR-1991

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00068
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: BO 37078

; TELEPHONE: 703/521-2297
; TELEFAX: 703/685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-129-151A-2

Query Match 95.7%; Score 291; DB 1; Length 57;
Best Local Similarity 96.5%; Pred. No. 1,3e-26;

Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVS 57
Db 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVS 57

RESULT 7
US-08-715-579-2

Sequence 2, Application US/08715579
Patent No. 5928946
GENERAL INFORMATION:
APPLICANT: DE VOS, Willem M.
APPLICANT: SIEZEN, Roelant J.
APPLICANT: KUIPERS, Oscar P.
TITLE OF INVENTION: LANTIBIOTICS SIMILAR TO NISIN A, LACTIC
TITLE OF INVENTION: ACID BACTERIA WHICH PRODUCE SUCH LANTIBIOTICS, METHOD FOR
TITLE OF INVENTION: CONSTRUCTING SUCH LACTIC ACID BACTERIA AND METHOD FOR
TITLE OF INVENTION: PRESERVING FOODSTUFFS WITH ETC.
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: C/O YOUNG & THOMPSON
STREET: 745 South 23rd Street, Second Floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/715,579
FILING DATE: 19-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/129,151
FILING DATE: 07-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NL 9100634
FILING DATE: 11-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00068
FILING DATE: 09-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: BO 37078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703/521-2297
TELEFAX: 703/685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-715-579-2

Query Match 95.7%; Score 291; DB 2; Length 57;
Best Local Similarity 96.5%; Pred. No. 1.3e-28;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVS 57
Db 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVS 57

RESULT 8
US-08-836-687B-40

Sequence 40, Application US/08836687B
Patent No. 6448034
GENERAL INFORMATION:
APPLICANT: Gasson, Michael John
APPLICANT: Dodd, Helen Mair
TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
FILE REFERENCE: 20747/70
CURRENT APPLICATION NUMBER: US/08/836,687B
CURRENT FILING DATE: 1995-11-20
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 40
LENGTH: 57
TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-40

Query Match 95.7%; Score 291; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 1.3e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVS 57
Db 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVS 57

RESULT 9

US-08-836-687B-43
Sequence 43, Application US/08836687B
Patent No. 6448034
GENERAL INFORMATION:
APPLICANT: Gasson, Michael John
APPLICANT: Dodd, Helen Mair
TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
FILE REFERENCE: 20747/70
CURRENT APPLICATION NUMBER: US/08/836,687B
CURRENT FILING DATE: 1995-11-20
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 43
LENGTH: 57
TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-43

Query Match 95.7%; Score 291; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 1.3e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVS 57
Db 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVS 57

RESULT 10

US-08-836-687B-42
Sequence 42, Application US/08836687B
Patent No. 6448034
GENERAL INFORMATION:
APPLICANT: Gasson, Michael John
APPLICANT: Dodd, Helen Mair
TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
FILE REFERENCE: 20747/70
CURRENT APPLICATION NUMBER: US/08/836,687B
CURRENT FILING DATE: 1995-11-20
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 42
LENGTH: 57
TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-42

Query Match 95.1%; Score 289; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 2, 2e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATCCHSIHVK 57
Db 1 MSTKDFNLDVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATCCHSIHVK 57

RESULT 11

US-08-836-687B-46
; Sequence 46, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 46
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-46

Query Match 95.1%; Score 289; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 2, 2e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATCCHSIHVK 57
Db 1 MSTKDFNLDVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATCCHSIHVK 57

RESULT 12

US-08-836-687B-44
; Sequence 44, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 44
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-44

Query Match 94.7%; Score 288; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 2, 9e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATCCHSIHVK 57
Db 1 MSTKDFNLDVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATCCHSIHVK 57

RESULT 13

US-08-836-687B-45
; Sequence 45, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN

; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 45
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.

US-08-836-687B-45

Query Match 93.8%; Score 285; DB 4; Length 57;
Best Local Similarity 94.7%; Pred. No. 6, 8e-28;
Matches 54; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATCCHSIHVK 57
Db 1 MSTKDFNLDVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATCCHSIHVK 57

RESULT 14

US-08-773-731A-2
; Sequence 2, Application US/08773731A
; Patent No. 6100056
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael J.
; APPLICANT: Dodd, Helen M.
; TITLE OF INVENTION: NISINS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESS: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP
; STREET: 2101 L Street N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/773,731A
; FILING DATE: 24-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/313,123
; FILING DATE: 18-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB93/00676
; FILING DATE: 01-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9207267.7
; FILING DATE: 02-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, Jr., James W.
; REGISTRATION NUMBER: 32,115
; REFERENCE/DOCKET NUMBER: E8280, 016/P016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-785-9700
; TELEFAX: 202-887-0689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-773-731A-2

Query Match 93.8%; Score 285; DB 3; Length 83;
Best Local Similarity 94.7%; Pred. No. 1e-27;
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSTKDENVLDVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
 |||||
Dd 1 MSTKDENVLDVSVSKDSGASPRITSIXLCTPGCKTGAIGMCMKMTATCXCSIHVSK 57

RESULT 15
ITE-09-036

US-08-836-687B-41
 ? Sequence 41: Application US/08836687B
 ? Patent No. 6448034
 ? GENERAL INFORMATION:
 ? APPLICANT: Gasson, Michael John
 ? APPLICANT: Dodd, Helen Mair
 ? TITLE OF INVENTION: PRODUCTION OF VARIANT
 ? FILE REFERENCE: 20747/70
 ? CURRENT APPLICATION NUMBER: US/08/836,687B
 ? CURRENT FILING DATE: 1995-11-20
 ? NUMBER OF SEQ ID NOS: 51
 ? SOFTWARE: PatentIn Ver. 2.1
 ? SEQ ID NO 41
 ? LENGTH: 56
 ? TYPE: PRF
 ? ORGANISM: *Lactococcus* sp.
 US-08-836-687B-41

Query Match	92.9%;	Score 282.5;	DB 4;	Length 56;
Best Local Similarity	96.5%;	Pred. No. 1.3e-27;		
Matches 55; Conservative	0;	Mismatches 1;	Indels 1;	Gaps 1;

QY 1 MSTKDENLDLVSYSKDSGASPRITSTSLCTPGCKTGALWGMCKMKTATCHCSIHVSK 57
|||||
Db 1 MSTKDFNLDLVSYSKDSGASPRITSLCTPGCKTGALWGMCKN-KTATCHCSIHVSK 56
|||||

Search completed: January 12, 2004, 14:18:03
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 14:03:04 ; Search time 47 Seconds
(without alignments)
244.384 Million cell updates/sec

Title: US-10-082-618-5
Perfect score: 304
Sequence: 1 MSTQDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCKMKATKCHSIHVK 57

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues
Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	/cgnt2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgnt2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgnt2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgnt2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgnt2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgnt2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgnt2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgnt2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgnt2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgnt2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgnt2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgnt2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgnt2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgnt2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgnt2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgnt2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17:	/cgnt2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18:	/cgnt2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	304	100.0	57	12	US-10-082-618-5	Sequence 5, Appli
2	298	98.0	57	12	US-10-372-098-8	Sequence 8, Appli
3	185	60.9	34	9	US-09-030-619-206	Sequence 206, App
4	185	60.9	34	9	US-09-917-340-14	Sequence 14, Appli
5	185	60.9	34	12	US-10-277-233-206	Sequence 206, App
6	157	51.6	56	9	US-09-030-619-208	Sequence 208, App
7	157	51.6	56	12	US-10-372-098-9	Sequence 9, Appli
8	157	51.6	56	12	US-10-082-618-4	Sequence 4, Appli
9	157	51.6	56	12	US-10-277-233-208	Sequence 208, App
10	94.5	31.1	52	12	US-10-372-098-10	Sequence 10, Appli
11	94.5	31.1	52	12	US-10-082-618-6	Sequence 6, Appli
12	91.5	30.1	52	12	US-10-082-618-1	Sequence 1, Appli
13	73	24.0	708	15	US-10-184-644-211	Sequence 211, App
14	73	24.0	708	15	US-10-184-634-211	Sequence 211, App
15	70	23.0	4185	12	US-10-140-472-67	Sequence 67, Appli

ALIGNMENTS

16	70	23.0	4185	12	US-10-141-761-67	Sequence 67, Appli
17	70	23.0	4185	12	US-10-142-885-67	Sequence 67, Appli
18	70	23.0	4185	12	US-10-158-790-67	Sequence 67, Appli
19	70	23.0	4185	12	US-10-137-871-67	Sequence 67, Appli
20	70	23.0	4185	12	US-10-140-805-67	Sequence 67, Appli
21	70	23.0	4185	12	US-10-140-864-67	Sequence 67, Appli
22	70	23.0	4185	12	US-10-140-923-67	Sequence 67, Appli
23	70	23.0	4185	12	US-10-141-756-67	Sequence 67, Appli
24	70	23.0	4185	12	US-10-141-759-67	Sequence 67, Appli
25	70	23.0	4185	15	US-10-123-155-67	Sequence 67, Appli
26	70	23.0	4185	16	US-10-146-731-67	Sequence 67, Appli
27	68.5	22.5	1570	15	US-10-184-644-335	Sequence 335, App
28	68.5	22.5	1570	15	US-10-184-634-335	Sequence 335, App
29	68.5	22.5	1617	12	US-10-063-685-61	Sequence 61, Appli
30	68.5	22.5	1617	15	US-10-184-644-235	Sequence 235, App
31	68.5	22.5	1617	15	US-10-184-634-235	Sequence 235, App
32	68.5	22.5	1875	12	US-10-140-472-317	Sequence 317, App
33	68.5	22.5	1875	12	US-10-141-761-317	Sequence 317, App
34	68.5	22.5	1875	12	US-10-142-885-317	Sequence 317, App
35	68.5	22.5	1875	12	US-10-158-790-317	Sequence 317, App
36	68.5	22.5	1875	12	US-10-137-871-317	Sequence 317, App
37	68.5	22.5	1875	12	US-10-140-805-317	Sequence 317, App
38	68.5	22.5	1875	12	US-10-140-864-317	Sequence 317, App
39	68.5	22.5	1875	12	US-10-140-923-317	Sequence 317, App
40	68.5	22.5	1875	12	US-10-141-756-317	Sequence 317, App
41	68.5	22.5	1875	12	US-10-141-759-317	Sequence 317, App
42	68.5	22.5	1875	15	US-10-123-155-317	Sequence 317, App
43	68.5	22.5	1875	16	US-10-146-731-317	Sequence 317, App
44	68	22.4	1660	12	US-10-063-685-31	Sequence 31, Appli
45	68	22.4	1660	15	US-10-184-644-147	Sequence 147, App

RESULT 1

US-10-082-618-5

Sequence 5, Application US/10082618

Publication No. US20030175207A1

GENERAL INFORMATION:

APPLICANT: OLSTEIN, ALAN D.

TITLE OF INVENTION: BACTERIOICIN-METAL COMPLEXES IN THE DETECTION OF

TITLE OF INVENTION: PATROGENS AND OTHER BIOLOGICAL ANALYTES

FILE REFERENCE: 7005-0003

CURRENT APPLICATION NUMBER: US/10/082,618

CURRENT FILING DATE: 2002-02-22

NUMBER OF SEQ ID NOS: 8

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5

LENGTH: 57

TYPE: PRT

ORGANISM: Lactococcus lactis

US-10-082-618-5

Query Match

Best Local Similarity 100.0%; Score 304; DB 12; Length 57;

Pred. No. 2.5e-29;

Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB

1 MSTQDFNLDVSVSKDGSASPRITSTSLCTPGCKTGALMGCKMKATKCHSIHVK 57

RESULT 2

US-10-372-098-8

Sequence 8, Application US/10372098

Publication No. US20030166835A1

GENERAL INFORMATION:

APPLICANT: UNIVERSITY OF MARYLAND

TITLE OF INVENTION: SUBLANTIN ANTIBIOTIC PRODUCED BY BACILLUS SUBTILIS 168

FILE REFERENCE: 8172-8072

CURRENT APPLICATION NUMBER: US/10/372,098

;; CURRENT FILING DATE: 2003-02-25
;; PRIOR APPLICATION NUMBER: US/09/462,478A
;; PRIOR FILING DATE: 2000-04-17
;; PRIOR APPLICATION NUMBER: US/60/053,035
;; PRIOR FILING DATE: 1997-07-18
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 8
;; LENGTH: 57
;; TYPE: PRT
;; ORGANISM: Unknown
;; FEATURE:
;; OTHER INFORMATION: Nisin A
US-10-372-098-8

Query Match 98.0%; Score 298; DB 12; Length 57;
Best Local Similarity 98.2%; Pred. No. 1.3e-28;
Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTKDPLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 MSTKDPLVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 3
US-09-030-619-206
;; Sequence 206, Application US/09030619B
;; Patent No. US20020035061A1
;; GENERAL INFORMATION:
;; APPLICANT: Krieger, Timothy J.
;; APPLICANT: Taylor, Robert
;; APPLICANT: Erfle, Douglas
;; APPLICANT: Frazer, Janet R.
;; APPLICANT: West, Michael H.P.
;; APPLICANT: McNicol, Patricia J.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
;; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
;; FILE REFERENCE: 660081.406
;; CURRENT APPLICATION NUMBER: US/09/030,619B
;; CURRENT FILING DATE: 1998-02-25
;; NUMBER OF SEQ ID NOS: 232
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 206
;; LENGTH: 34
;; TYPE: PRT
;; ORGANISM: Lactococcus lactis
US-09-030-619-206

Query Match 60.9%; Score 185; DB 9; Length 34;
Best Local Similarity 97.1%; Pred. No. 2.8e-15;
Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 24 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 34

RESULT 4
US-09-917-340-14
;; Sequence 14, Application US/09917340
;; Patent No. US20020090369A1
;; GENERAL INFORMATION:
;; APPLICANT: Murphy, Christopher J.
;; APPLICANT: McNulty, Jonathan F.
;; APPLICANT: Reid, Ted W.
;; TITLE OF INVENTION: Transplant Media
;; FILE REFERENCE: TPLANT-06468
;; CURRENT APPLICATION NUMBER: US/09/917,340
;; CURRENT FILING DATE: 2001-07-29
;; PRIOR APPLICATION NUMBER: 60/221,632
;; PRIOR FILING DATE: 2000-07-28
;; PRIOR APPLICATION NUMBER: 60/249,602

;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/290,932
;; PRIOR FILING DATE: 2001-05-15
;; NUMBER OF SEQ ID NOS: 96
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 14
;; LENGTH: 34
;; TYPE: PRT
;; ORGANISM: Lactococcus lactis
US-09-917-340-14

Query Match 60.9%; Score 185; DB 9; Length 34;
Best Local Similarity 97.1%; Pred. No. 2.8e-15;
Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 24 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 34

RESULT 5
US-10-277-233-206
;; Sequence 206, Application US/10277233
;; Publication No. US20030232750A1
;; GENERAL INFORMATION:
;; APPLICANT: Krieger, Timothy J.
;; APPLICANT: Taylor, Robert
;; APPLICANT: Erfle, Douglas
;; APPLICANT: Frazer, Janet R.
;; APPLICANT: West, Michael H.P.
;; APPLICANT: McNicol, Patricia J.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
;; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
;; FILE REFERENCE: 660081.406C1
;; CURRENT APPLICATION NUMBER: US/10/277,233
;; CURRENT FILING DATE: 2002-10-18
;; NUMBER OF SEQ ID NOS: 232
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 206
;; LENGTH: 34
;; TYPE: PRT
;; ORGANISM: Lactococcus lactis
US-10-277-233-206

Query Match 60.9%; Score 185; DB 12; Length 34;
Best Local Similarity 97.1%; Pred. No. 2.8e-15;
Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 24 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
DB 1 ITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 34

RESULT 6
US-09-030-619-208
;; Sequence 208, Application US/09030619B
;; Patent No. US20020035061A1
;; GENERAL INFORMATION:
;; APPLICANT: Krieger, Timothy J.
;; APPLICANT: Taylor, Robert
;; APPLICANT: Erfle, Douglas
;; APPLICANT: Frazer, Janet R.
;; APPLICANT: West, Michael H.P.
;; APPLICANT: McNicol, Patricia J.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
;; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
;; FILE REFERENCE: 660081.406
;; CURRENT APPLICATION NUMBER: US/09/030,619B
;; CURRENT FILING DATE: 1998-02-25
;; NUMBER OF SEQ ID NOS: 232
;; SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 208
LENGTH: 56
TYPE: PRT
ORGANISM: Bacillus subtilis
US-09-030-619-208

Query Match 51.6%; Score 157; DB 9; Length 56;
Best Local Similarity 59.2%; Pred. No. 1,1e-11;
Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 5 DFNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 53
DB 6 DFDLDVVVKVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 54

RESULT 7
US-10-372-098-9

Sequence 9, Application US/10372098
Publication No. US20030166835A1
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBSTANTIAL INHIBITION PRODUCED BY BACILLUS SUBTILIS 168
FILE REFERENCE: 8172-8072
CURRENT APPLICATION NUMBER: US/10/372,098
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: US/09/462,478A
PRIOR FILING DATE: 2000-04-17
PRIOR APPLICATION NUMBER: US/60/053,035
PRIOR FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9

LENGTH: 56
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: subtilin
US-10-372-098-9

Query Match 51.6%; Score 157; DB 12; Length 56;
Best Local Similarity 59.2%; Pred. No. 1,1e-11;
Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 5 DFNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 53
DB 6 DFDLDVVVKVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 54

RESULT 8
US-10-082-618-4

Sequence 4, Application US/10082618
Publication No. US20030175207A1
GENERAL INFORMATION:
APPLICANT: OLSTEIN, ALAN D.
APPLICANT: FEIRTAG, JOELEN
TITLE OF INVENTION: BACTERIOCIDAL METAL COMPLEXES IN THE DETECTION OF
TITLE OF INVENTION: PATHOGENS AND OTHER BIOLOGICAL ANALYTES
FILE REFERENCE: 7005-0003
CURRENT APPLICATION NUMBER: US/10/082,618
CURRENT FILING DATE: 2002-02-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 56
TYPE: PRT
ORGANISM: Bacillus subtilis
US-10-082-618-4

Query Match 51.6%; Score 157; DB 12; Length 56;
Best Local Similarity 59.2%; Pred. No. 1,1e-11;
Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 5 DFNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 53

DB 6 DFDLDVVVKVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 54

RESULT 9
US-10-277-233-208

Sequence 208, Application US/10277233
Publication No. US20030232750A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 66081.406C1
CURRENT APPLICATION NUMBER: US/10/277,233
CURRENT FILING DATE: 2002-10-18
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 208
LENGTH: 56
TYPE: PRT
ORGANISM: Bacillus subtilis
US-10-277-233-208

Query Match 51.6%; Score 157; DB 12; Length 56;
Best Local Similarity 59.2%; Pred. No. 1,1e-11;
Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 5 DFNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 53
DB 6 DFDLDVVVKVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 54

RESULT 10
US-10-372-098-10

Sequence 10, Application US/10372098
Publication No. US20030166835A1
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBSTANTIAL INHIBITION PRODUCED BY BACILLUS SUBTILIS 168
FILE REFERENCE: 8172-8072
CURRENT APPLICATION NUMBER: US/10/372,098
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: US/09/462,478A
PRIOR FILING DATE: 2000-04-17
PRIOR APPLICATION NUMBER: US/60/053,035
PRIOR FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 52
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: epidermin
US-10-372-098-10

Query Match 31.1%; Score 94.5; DB 12; Length 52;
Best Local Similarity 62.2%; Pred. No. 0.00033;
Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;

QY 6 FNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 38
DB 11 FNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTATGCSI 46

RESULT 11
US-10-082-618-6

```

; Sequence 6, Application US/10082618
; Publication No. US20030175207A1
; GENERAL INFORMATION:
; APPLICANT: OLSTEIN, ALAN D.
; APPLICANT: PEIRTAG, JOELEN
; TITLE OF INVENTION: BACTERIOCIN-METAL COMPLEXES IN THE DETECTION OF
; FILE OF INVENTION: PATHOGENS AND OTHER BIOLOGICAL ANALYTES
; FILE REFERENCE: 7005-0003
; CURRENT APPLICATION NUMBER: US/10/082,618
; CURRENT FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 52
; TYPE: PRF
; ORGANISM: Staphylococcus epidermidis
US-10-082-618-6

Query Match      31.1%; Score 94.5; DB 12; Length 52;
Best Local Similarity 62.2%; Pred. No. 0.00033;
Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;

QY      6 FNLDVSVSK--DSGASPRITSLCTPGC-KTGA 38
DB      11 FNLD-VKNVAKESNDSGAEPRISKFLCTPGCAKTGS 46

RESULT 12
US-10-082-618-1
; Sequence 1, Application US/10082618
; Publication No. US20030175207A1
; GENERAL INFORMATION:
; APPLICANT: OLSTEIN, ALAN D.
; APPLICANT: PEIRTAG, JOELEN
; TITLE OF INVENTION: BACTERIOCIN-METAL COMPLEXES IN THE DETECTION OF
; FILE OF INVENTION: PATHOGENS AND OTHER BIOLOGICAL ANALYTES
; FILE REFERENCE: 7005-0003
; CURRENT APPLICATION NUMBER: US/10/082,618
; CURRENT FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 52
; TYPE: PRF
; ORGANISM: Staphylococcus gallinarum
US-10-082-618-1

Query Match      30.1%; Score 91.5; DB 12; Length 52;
Best Local Similarity 62.2%; Pred. No. 0.00076;
Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;

QY      6 FNLDVSVSK--DSGASPRITSLCTPGC-KTGA 38
DB      11 FNLD-VKNVAKESNDSGAEPRISKFLCTPGCAKTGS 46

RESULT 13
US-10-184-644-211
; Sequence 211, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 211
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-211

Query Match      24.0%; Score 73; DB 15; Length 708;
Best Local Similarity 43.2%; Pred. No. 2.3;
Matches 16; Conservative 4; Mismatches 15; Indels 2; Gaps 1;

QY      18 SGASPRITSLCTPGCKTGALMGCMKAT--CHCS 52
DB      341 TGACCGCTTCTCTGCGATGACCGACCTGTGGCGCT 377

RESULT 14
US-10-184-634-211
; Sequence 211, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 211
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-211

Query Match      24.0%; Score 73; DB 15; Length 708;
Best Local Similarity 43.2%; Pred. No. 2.3;
Matches 16; Conservative 4; Mismatches 15; Indels 2; Gaps 1;

QY      18 SGASPRITSLCTPGCKTGALMGCMKAT--CHCS 52
DB      341 TGACCGCTTCTCTGCGATGACCGACCTGTGGCGCT 377

RESULT 15
US-10-140-472-67
; Sequence 67, Application US/10140472
; Publication No. US20030138868A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
```

; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P330R1C168
 ; CURRENT APPLICATION NUMBER: US/10/140,472
 ; CURRENT FILING DATE: 2002-05-06
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 67
 ; LENGTH: 4185
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-140-472-67

Query Match 23.0%; Score 70; DB 12; Length 4185;
 Best Local Similarity 46.4%; Pred. No. 38;
 Matches 13; Conservative 2; Mismatches 11; Indels 2; Gaps 1;
 QY 25 TSTSLCTPGCKTGALMGCMKMTATCHCS 52
 DB 33 TGTATCTTGGCTTG--GCTATCTTCCT 58

Search completed: January 12, 2004, 14:23:23
 Job time : 49 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 14:22:27 ; Search time 26 Seconds
(without alignments)
210.831 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304
Sequence: 1 MSTDQFNDLIVSVSKDGA.....ALMGCMKATCHGSIHVK 57

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	298	98.0	57	1	NILISA
2	157	51.6	36	1	subtilin precursor -
3	94.5	31.1	52	1	EPSED
4	91.5	30.1	52	1	EPSED
5	63	20.7	995	2	SS0358
6	59.5	19.6	78	2	C82577
7	59.5	19.6	330	2	T25169
8	59.5	19.6	1574	2	T13954
9	59	19.4	456	2	S20597
10	58.5	19.2	166	2	T33970
11	58.5	19.2	772	2	D56695
12	58	19.1	565	2	F84721
13	57.5	18.9	727	2	A56879
14	57.5	18.9	771	2	S35681
15	56.5	18.6	232	2	S70355
16	56.5	18.6	456	2	A31857
17	56.5	18.6	741	2	B49555
18	56.5	18.6	1700	2	S08167
19	56	18.4	107	2	T49527
20	56	18.4	506	2	AT3300
21	56	18.4	1016	2	T41720
22	55.5	18.3	333	2	T29208
23	55.5	18.3	724	2	T47149
24	55.5	18.3	1620	2	T37283
25	55.5	18.3	1891	2	T13594
26	55.5	18.3	1920	2	T13893
27	55	18.1	267	2	C85075
28	55	18.1	369	2	S13721
29	55	18.1	369	2	S13721

30	55	18.1	374	1	A56436	alcohol dehydrogen
31	55	18.1	384	2	S25771	gaal protein - mou
32	55	18.1	551	2	S23400	bud emergence medi
33	55	18.1	782	2	A10062	conserved hypothet
34	55	18.1	808	1	S62594	replication licens
35	54.5	17.9	134	2	T22275	hypothetical prote
36	54	17.8	172	2	T50694	transcription fact
37	54	17.8	199	2	T47716	transcription fact
38	54	17.8	343	2	T30233	ornithine cyclodea
39	54	17.8	355	2	C96551	protein T3P18.9 (1
40	54	17.8	511	2	T07787	pyruvate kinase (E
41	54	17.8	518	2	S42091	Tid(56) protein -
42	54	17.8	876	2	B82163	DNA topoisomerase
43	54	17.8	1030	2	H96568	unknown protein, 2
44	54	17.8	1046	2	AD2959	AcB/AcrB/AcrF fam
45	54	17.8	1046	2	B98324	probable rnd efflu

ALIGNMENTS

RESULT 1
NILISA
nlsin precursor - Lactococcus lactis
N/Alternate names: nlsin A; nlsin Z
C/Species: Lactococcus lactis
C/Date: 21-May-1990 #sequence revision 12-May-1994 #text change 18-Jun-1999
C/Accession: A31915; A32809; B45821; A43743; S17856; B48951; S36734; S70485; S16779; S36;
R/Buchanan, G.W.; Banerjee, S.; Hansen, J.N.
J. Biol. Chem. 263, 16260-16266, 1988
A/Title: Structure, expression, and evolution of a gene encoding the precursor of nlsin,
A/Reference number: A92679; MUID:89034093; PMID:3141403
A/Accession: A31915
A/Molecule type: DNA
A/Residues: 1-57 <BUC>
A/Cross-references: GB:704057; NID:G153816; PIDN:AAA8606.1; PID:G153817
A/Experimental source: ATCC 11454
A/Note: The authors identified the species as Streptococcus lactis
R/Kaletta, C.; Entian, K.D.
J. Bacteriol. 171, 1597-1601, 1989
A/Title: Nlsin, a peptide antibiotic: cloning and sequencing of the nlsA gene and postre
A/Reference number: A32809; MUID:89155467; PMID:2493449
A/Accession: A32809
A/Molecule type: DNA
A/Residues: 1-57 <KAL>
A/Cross-references: GB:M24527; NID:G341189; PIDN:AAA26948.1; PID:G530218
R/Dodd, H.M.; Horn, N.; Gasson, M.J.
J. Gen. Microbiol. 136, 555-566, 1990
A/Title: Analysis of the genetic determinant for production of the peptide antibiotic nls
A/Reference number: A45821; MUID:90362041; PMID:2118169
A/Accession: B45821
A/Molecule type: DNA
A/Residues: 1-57 <DOD>
A/Cross-references: GB:M27277; NID:G149440; PIDN:AAA25188.1; PID:G149441
R/Siezen, M.T.; Chung, Y.J.; Hansen, J.N.
Appl. Environ. Microbiol. 57, 1181-1188, 1991
A/Title: Characterization of the nlsin gene as part of a polycistronic operon in the chrC
A/Reference number: A43743; MUID:91282469; PMID:1905517
A/Accession: A43743
A/Molecule type: DNA
A/Residues: 1-57 <STE>
A/Cross-references: GB:M65059; NID:G149447; PIDN:AAA73038.1; PID:G149448
A/Experimental source: ATCC 11454
R/Mulders, J.W.M.; Boerigter, I.; Rollema, H.S.; Siezen, R.J.; de Vos, W.M.
Eur. J. Biochem. 201, 581-584, 1991
A/Title: Identification and characterization of the lactobiotic nlsin Z, a natural nlsin
A/Reference number: S17858; MUID:9203612; PMID:1995953
A/Accession: S17858
A/Molecule type: DNA
A/Residues: 1-49, 'N', 51-57 <MOD>
A/Cross-references: EMBL:X61144; NID:G44046; PIDN:CMA43440.1; PID:G44047
A/Experimental source: strain NIZO 22186
A/Note: nlsin Z allelic variant; amino acid composition and structure determination by N

R.Engelke, G.; Gutowski-Eckel, Z.; Hammelmann, M.; Entian, K.D.
 Appl. Environ. Microbiol. 58, 3730-3743, 1992
 A>Title: Biosynthesis of the lantibiotic nisin: genomic organization and membrane local
 A>Reference number: A48951; MUID:933128945; PMID:1482192
 A>Accession: B48951
 A>Molecule type: DNA
 A>Residues: 1-57 <ENG>
 A>Cross-references: GB:K6307; NID:944040; PIDN:CAA6380.1; PID:944042
 A>Experimental source: strain 6P3
 A>Note: sequence extracted from NCBI backbone (NCBIN:122292, NCBI:122295)
 R>Kuipers, O.P.; Beershuizen, M.M.; Stezen, R.J.; de Vos, W.M.
 Eur. J. Biochem. 216, 281-291, 1993
 A>Title: Characterization of the nisin gene cluster nlsABTCIPR of *Lactococcus lactis*. R
 A>Reference number: S36734; MUID:93373937; PMID:7689965
 A>Accession: S36734
 A>Status: preliminary
 A>Molecule type: DNA
 A>Residues: 1-57 <KU>
 A>Cross-references: GB:L16226; NID:9400365; PIDN:AAA25189.1; PID:9400366
 R>Gross, E.; Morell, J.L.
 J. Am. Chem. Soc. 93, 4634-4635, 1971
 A>Title: The structure of nisin
 A>Reference number: A54460; MUID:72072901; PMID:5131162
 A>Contents: annotation
 R>Kuipers, O.P.; Rollema, H.S.; de Vos, W.M.; Stezen, R.J.
 FEBS Lett. 330, 23-27, 1993
 A>Title: Biosynthesis and secretion of a precursor of nisin Z by *Lactococcus lactis*, dir
 A>Reference number: S36142; MUID:93380562; PMID:8370453
 A>Contents: annotation
 R>van der Meer, J.R.; Polman, J.; Beershuizen, M.M.; Stezen, R.J.; Kuipers, O.P.; De Vos
 J. Bacteriol. 175, 2578-2588, 1993
 A>Title: Characterization of the *Lactococcus lactis* nisin A operon genes nlsP, encoding
 lived in nisin biosynthesis.
 A>Reference number: A40621; MUID:93239683; PMID:8478324
 A>Contents: annotation
 R>Chan, W.C.; Leyland, M.; Clark, J.; Dodd, H.M.; Lian, L.Y.; Gasson, M.J.; Bycroft, B.W
 FEBS Lett. 390, 129-132, 1996
 A>Title: Structure-activity relationships in the peptide antibiotic nisin: antibacterial
 A>Reference number: S70485; MUID:96505786; PMID:8706842
 A>Accession: S70485
 A>Molecule type: protein
 A>Residues: 24-57 <CHA>
 C>Comment: Nisin is secreted as an inactive precursor and then activated by cleavage with
 C>Genetics:
 A>Gene: *spas*, *nlsA*
 C>Superfamily: subtilin precursor
 C>Keywords: antibiotic; lantibiotic
 F:2-23/Domains: propeptide #status experimental <PRO>
 F:24-57/Product: nisin #status experimental <MAT>
 F:23-24/Cleavage site: Arg-11e (protease NlsP) #status experimental
 F:25/Modified site: dehydrobutyryne (Thr) #status experimental
 F:26-30/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
 F:28/Modified site: dehydroalanine (Ser) #status experimental
 F:31-34/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:36-42/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:46-49/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:48-51/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:56/Modified site: dehydroalanine (Ser) #status experimental

C>Species: *Bacillus subtilis*
 C>Date: 30-Sep-1989 #sequence revision 12-May-1994 #text change 21-Jul-2000
 C>Accession: A28112; D42655; D43935; I40514; I39980; S36142
 R>Banerjee, S.; Hansen, J.N.
 J. Biol. Chem. 263, 9508-9514, 1988
 A>Title: Structure and expression of a gene encoding the precursor of subtilin, a small I
 A>Reference number: A28112; MUID:88243844; PMID:2837490
 A>Accession: A28112
 A>Molecule type: DNA
 A>Residues: 1-56 <BAN>
 A>Cross-references: GB:U03767; NID:9143718; PIDN:AAA22841.1; PID:9143719
 R>Chung, Y.J.; Steen, M.T.; Hansen, J.N.
 J. Bacteriol. 174, 1417-1422, 1992
 A>Title: The subtilin gene of *Bacillus subtilis* ATCC 6633 is encoded in an operon that cc
 A>Reference number: A42655; MUID:92138640; PMID:1735728
 A>Accession: D42655
 A>Molecule type: DNA
 A>Residues: 1-56 <CHU>
 A>Cross-references: GB:M83944; NID:9143557; PIDN:AAA22772.1; PID:9143561
 A>Experimental source: ATCC 6633
 A>Note: sequence extracted from NCBI backbone (NCBI:79670)
 R>Klein, C.; Kaletta, C.; Schmeil, N.; Entian, K.D.
 Appl. Environ. Microbiol. 58, 132-142, 1992
 A>Title: Analysis of genes involved in biosynthesis of the lantibiotic subtilin.
 A>Reference number: A43935; MUID:92171481; PMID:1539969
 A>Accession: D43935
 A>Molecule type: DNA
 A>Residues: 1-56 <KLE>
 A>Cross-references: GB:M86869; NID:9143713; PIDN:AAA22840.1; PID:9143717
 A>Experimental source: ATCC 6633
 A>Note: sequence extracted from NCBI backbone (NCBI:84011, NCBI:84020)
 R>Klein, C.; Entian, K.D.
 Appl. Environ. Microbiol. 60, 2793-2801, 1994
 A>Title: Genes involved in self-protection against the lantibiotic subtilin produced by E
 A>Reference number: I40511; MUID:94368094; PMID:8085823
 A>Accession: I40511
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A>Molecule type: DNA
 A>Residues: 1-56 <KLE>
 A>Cross-references: EMBL:U09819; NID:92702240; PIDN:AAA91589.1; PID:9595319
 A>Experimental source: ATCC 6633
 R>Chan, W.C.; Bycroft, B.W.; Leyland, M.T.; Lian, L.Y.; Yang, J.C.; Roberts, G.C.K.
 FEBS Lett. 300, 56-62, 1992
 A>Title: Sequence-specific resonance assignment and conformational analysis of subtilin I
 A>Reference number: A44571; MUID:92192284; PMID:1547888
 A>Contents: annotation
 R>Chan, W.C.; Bycroft, B.W.; Leyland, M.T.; Lian, L.Y.; Roberts, G.C.K.
 Biochem. J. 291, 23-27, 1993
 A>Title: A novel post-translational modification of the peptide antibiotic subtilin: isoI
 A>Reference number: A53265; MUID:9322611; PMID:8471040
 A>Contents: annotation; amino-terminal succinylation; ATCC 6633
 R>Kuipers, O.P.; Rollema, H.S.; de Vos, W.M.; Stezen, R.J.
 FEBS Lett. 330, 23-27, 1993
 A>Title: Biosynthesis and secretion of a precursor of nisin Z by *Lactococcus lactis*, dire
 A>Reference number: S36142; MUID:93380562; PMID:8370453
 A>Contents: annotation
 C>Comment: *Bacillus subtilis* strain ATCC 6633 carries this gene that is not found in stre
 C>Genetics:
 A>Gene: *spas*
 C>Superfamily: subtilin precursor
 C>Keywords: antibiotic; blocked amino end; lantibiotic
 F:2-24/Domains: propeptide #status experimental <SIG>
 F:25-56/Product: subtilin A #status experimental <MAT>
 F:25-56/Product: subtilin B #status experimental <MAT>
 F:25/Modified site: succinylated amino end (Tyr) (in mature form) #link MATB #status expe
 F:27-31/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
 F:29/Modified site: dehydroalanine (Ser) #status experimental
 F:32-35/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:37-43/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:42/Modified site: (2)-dehydrobutyryne (Thr) #status experimental
 F:47-50/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:49-52/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:55/Modified site: dehydroalanine (Ser) #status experimental

Query Match 51.6%; Score 157; DB 1; Length 56;
Best Local Similarity 59.2%; Pred. No. 5.5e-11;
Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

OY 5 DPNLDVSVSKDGSASPRITSTSLCTPGCTGALMGCMKATATGCSI 53
DB 6 DFDLDVAVKSDSKITPQWSESICTPGCVGALQTCFLQTLTNCCKI 54

RESULT 3

epidermin precursor - Staphylococcus epidermidis plasmids

C:Species: Staphylococcus epidermidis
C:Date: 31-Dec-1988 #sequence_revision 12-May-1994 #text_change 18-Jun-1999

C:Accession: S00768; S23415; A61287
R:Schneil, N.; Entlan, K.D.; Schneider, U.; Goetz, F.; Zaehner, H.; Kellner, R.; Jung, G.
Nature 333, 276-278, 1988

A:Title: Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with fou
A:Reference number: S00768; MUID:88216821; PMID:2835685

A:Accession: S00768

A:Molecule type: DNA

A:Residues: 1-52 <SCH>
A:Cross-references: EMBL:X07840; NID:g46961; PIDN:CAA30689.1; PID:g46962

A:Note: plasmid pEP132
R:Schneil, N.; Engelke, G.; Augustin, J.; Rosenstein, R.; Ungertmann, V.; Goetz, F.; Entl
Eur. J. Biochem. 204, 57-68, 1992

A:Title: Analysis of genes involved in the biosynthesis of lantibiotic epidermin.
A:Reference number: S23413; MUID:92155237; PMID:1740156

A:Accession: S23413

A:Molecule type: DNA

A:Residues: 1-52 <SC2>
A:Cross-references: EMBL:X62386; NID:g46964; PIDN:CAA44252.1; PID:g46967

A:Note: plasmid pHe32
R:Allgater, H.; Jung, G.; Werner, R.G.; Schneider, U.; Zaehner, H.
Eur. J. Biochem. 160, 9-22, 1986

A:Title: Epidermin: sequencing of a heterodet tetracyclic 21-peptide amide antibiotic.
A:Reference number: A61287; MUID:87030262; PMID:3769923

A:Accession: A61287

A:Molecule type: protein

A:Residues: 31-52 <ALV>
A:Note: plasmid Tue3298

C:Genetics:

A:Gene: epia

A:Genome: plasmid

C:Superfamily: subtilin precursor

C:Keywords: antibiotic; blocked carboxyl end; lantionine

F:1-10/Domain: propeptide #status predicted <PRO>
F:31-52/Product: epidermin #status experimental <MAT>

F:33-37/Cross-link: sn-(2S,6R)-3-methyl-lantionine (Ser-Cys) #status experimental

F:38-41/Cross-link: (2S,3S,6R)-3-methyl-lantionine (Thr-Cys) #status experimental

F:44/Modified site: (2)-dehydrobutyryl (Thr) #status experimental

F:46-51/Cross-link: sn-(2S,6R)-lantionine (Ser-Cys) #status experimental

F:49-52/Cross-link: (S,Z)-S-(2-aminovinyl) cysteine (Ser-Cys) #status experimental

Query Match 31.1%; Score 94.5; DB 1; Length 52;
Best Local Similarity 62.2%; Pred. No. 0.00059;
Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;

OY 6 FNLDVSVSK---DSGASPRITSTSLCTPGC-KTGA 38
DB 11 FNLDVAVKSDSKITPQWSESICTPGCVGALQTCFLQTLTNCCKI 54

RESULT 4

epidermin precursor - Staphylococcus gallinarum

C:Species: Staphylococcus gallinarum

C:Date: 10-Mar-1994 #sequence_revision 12-May-1994 #text_change 26-Feb-1999

C:Accession: A61072; A44573; A53264

R:Schneil, N.; Entlan, K.D.; Goetz, F.; Hoerner, T.; Kellner, R.; Jung, G.
FEMS Microbiol. Lett. 58, 263-268, 1989

A:Title: Structural gene isolation and prepeptide sequence of gallidermin, a new lantibid

A:Reference number: A61072
A:Accession: A61072
A:Status: not compared with conceptual translation
A:Molecule type: DNA

A:Residues: 1-52 <SCH>
A:Kellner, R.; Jung, G.; Hoerner, T.; Zaehner, H.; Schneil, N.; Entlan, K.D.; Goetz, F.
Eur. J. Biochem. 177, 53-59, 1988

A:Title: Gallidermin: a new lantionine-containing polypeptide antibiotic.
A:Reference number: A44573; MUID:89030695; PMID:3181159

A:Accession: A44573

A:Molecule type: protein

A:Residues: 31-52 <KEL>
C:Genetics:

A:Gene: gdmA

C:Superfamily: subtilin precursor

C:Keywords: antibiotic; blocked carboxyl end; lantionine

F:1-10/Domain: propeptide #status predicted <PRO>
F:31-52/Product: gallidermin #status experimental <MAT>

F:33-37/Cross-link: sn-(2S,6R)-3-methyl-lantionine (Ser-Cys) #status experimental

F:38-41/Cross-link: (2S,3S,6R)-3-methyl-lantionine (Thr-Cys) #status experimental

F:44/Modified site: dehydrobutyryl (Thr) #status experimental

F:46-51/Cross-link: sn-(2S,6R)-lantionine (Ser-Cys) #status experimental

F:49-52/Cross-link: (S,Z)-S-(2-aminovinyl) cysteine (Ser-Cys) #status experimental

Query Match 30.1%; Score 91.5; DB 1; Length 52;
Best Local Similarity 62.2%; Pred. No. 0.0013;
Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;

OY 6 FNLDVSVSK---DSGASPRITSTSLCTPGC-KTGA 38
DB 11 FNLDVAVKSDSKITPQWSESICTPGCVGALQTCFLQTLTNCCKI 54

RESULT 5

hypothetical protein Y1169c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein Y19402.07c
C:Species: Saccharomyces cerevisiae

C:Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 19-Apr-2002

C:Accession: S50358

R:Live, G.; Bowman, S.; Churcher, C.
submitted to the EMBL Data Library, December 1994

A:Reference number: S50349

A:Accession: S50358

A:Molecule type: DNA

A:Residues: 1-995 <LYE>
A:Cross-references: GB:Z47047; EMBL:Z46921; NID:g603997; PID:g604006; GSPDB:GN00009; MIP

C:Genetics:

A:Gene: MIP8:Y1169c

A:Cross-references: SGD:S0001431

A:Map position: 9L

Query Match 20.7%; Score 63; DB 2; Length 995;
Best Local Similarity 26.9%; Pred. No. 27;
Matches 21; Conservative 8; Mismatches 25; Indels 24; Gaps 2;

OY 3 TKDNLVSVSKDGSASPRITSTSLCTPGC---KTGALM----- 40
DB 676 TTDSNGVYITTTVPSCSTTAITTSCEDECHVSTSGAVTETVSSKSYTATVTHCD 735

OY 41 --GCMKTATGCSIHVS 56
DB 736 DNGCNTKTVISECKETS 753

RESULT 6

hypothetical protein XF2284 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: C82577

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C62577
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-78 <SIM>
A:Cross-references: GB:AE004040; GB:AE003849; NID:g9107437; PIDN:AA85083.1; GSPDB:GN001040
A:Experimental source: strain 945C
R:Simpon, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H
de-Melo, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fromm
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Klieger, J.E.; Kuramae, E.E.; Laigret
chado, M.A.; Madeira, E.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Oliveira, E.C.; Miyaki, C.Y.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sasaki
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
A:Genetics:
A:Gene: XF2284

Query Match	19.6%	Score 59.5	DB 2	Length 78
Best Local Similarity	41.2%	Pred. No. 7.4		
Matches	14	Conservative	5	Mismatches 14; Indels 1; Gaps 1
QY	23	RITSLCTPGCK-TGALMGCMKTKATGCHSHV	55	
DB	15	RPDAVSVLTGPGCKGATLLTRVRATVYCSKTHI	48	

RESULT 7
T25169
hypothetical protein T251.6 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T25169
R:Wilkinson, J.

submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T251c9
A:Status: preliminary; translated from GB/EMBL/DBU
A:Molecule type: DNA
A:Residues: 1-330 <WIL>
A:Cross-references: EMBL:Z81129; P1DN:CA03405.1; GSPDB:GN00023; CESP:T23F1.6
A:Experimental source: clone T23F1
C:Genetics:
A:Gene: CESP:T23F1.6
A:Map position: 5
A:Insertions: 16/3
C:Superfamily: gliadin

	Query Match	19.6%	Score 59.5;	DB 2;	Length 330;
	Best Local Similarity	27.9%	Pred. No. 26;		
	Matches	12;	Mismatches	17;	Gaps 11
Oy	23	RITSTSLCTPGCKTGALMGCN-----MKTATCHSHVVS	56		
	:	:::	: :		
Db	275	RTTASQCCAPACSTSCNCSNQPAOMACQPMNSCGGQQNNIS	317		

RESULT 8
T13954
MEGF6 protein - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 20-Sep-1999 #sequence_20-Sep-1999 #text_change 21-Jul-2000
C/Accession: T13954
R:Nakajima, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998

A;Title: Identification of h9qh-molecular-weight proteins with multiple EGF-like motifs
A;Reference number: Z14126; MUID:98360089; PMID:963030
A;Accession: T13954
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1574 <NAK>
A;Cross-references: EMBL:AB011532; NID:G3449293; PIDN:BA32462.1; PID:G3449294
A;Experimental source: strain Sprague-Dawley; brain
C;Genetics:
A;Gene: MEGR6

Query Match 19.6%; Score 59.5; DB 2; Length 1574;
Best Local Similarity 27.8%; Pred. No. 1e+02;
Matches 15; Conservative 3; Mismatches 9; Indels 27; Gaps 2.

```

QY      23  RITSTSLCPG-----CTGALMGCMKATATC 51
          ||| ||| ||| ||| ||| ||| ||| |||
          :|: ||| ||| ||| ||| ||| ||| |||
Db      757  RVTGECLECPGKTGEGDGCADCEGRWGLCQEI 808
          ||| ||| ||| ||| ||| ||| ||| |||
          :|: ||| ||| ||| ||| ||| ||| |||

```

RESULT 9

ribonuclease inhibitor - rat
C:Species: Rattus norvegicus (Norway rat)
D:Accession: U00896 Date Recd: 10 Nov 1995 About: Abn 21 Jan 2000

CiDate: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
CiAccession: S20597
R.Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.

A:Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution
A:Reference number: S20597; MUID:92162755; PMID:1536887

A;Molecule type: mRNA

A;Residues: 1-456 <KAW>

F280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
 A1Cross-references: EMBL:4652528; NID:G57670; PDB:CAA44388.1; PID:G57671
 C:Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
 F280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match	19.4%;	Score 59;	DB 2;	Length 456;
Best Local Similarity	34.0%;	Pred. No. 39;		
Matches 17; Conservative	0;	Mismatches 19;	Indels 6;	Gaps 3;

```

Qy      8 LDIVSVSKDGSAPRITSTSLCTPGCKTGAL--MGMNMKTATC--HCSI 53
      | | | | | | | | | | | | | | | | | | | | | |
Db      284 LSLAGNELKDEGA--QLLCSLLEPGQLSLMVKTCSTLAASCPHFCSV 311

```

RESULT 10
T33970

hypothetical protein F46E10.11 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

A;Description: The sequence of *C. elegans* cosmid F46E10 submitted to the EMBL Data Library, February 1999

A;Reference number: Z21446

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1166 <J0H>
A:Cross-references: EMBL:AF125955; PDB:AMD1471.1; GSPDB:GN00023; CESP:F46E10.11
A:Experimental source: strain Bristol N2; clone F46E10

A;Gene: CESP:F46E10.11
A;Map position: 5
A;Introns: 55/1

```

Query Match      19.2%  Score 58.5;  DB 2;  Length 166;
Best Local Similarity 31.7%  Pred. No. 19;
Matches 19;  Conservative 4;  Mismatches 16;  Indels 21;  Gaps 3;

07      9  DLVSVSKKDSG-ASPR-----ITSLSICTPGCKTGALMGNNKTCACHS 52
      |  |||:::  :  ||  :  |  ||  |  ||  ||  ||  |

```

Db 19 DTASVSQESTEVARRPKRQIYYLCGFNFNYLSLTPCNSGSGST-----CNCNTATCSTS 73

RESULT 11

D56695
transducin-like enhancer-of-split homolog TLE-3 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 26-May-2000
C:Accession: D56695
R:Stifani, S.; Blumweller, C.M.; Redhead, N.J.; Hill, R.E.; Artavanis-Tsakonas, S.
Nature Genet. 2, 119-127, 1992
A:Title: Human homologs of a Drosophila enhancer of split gene product define a novel fa
A:Reference number: A56695; MUID:93265135; PMID:1303260
A:Accession: D56695
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-772 <STL>
A:Cross-references: GB:M99438; NID:9307513; PIDN:AAA61194.1; PID:9307514
C:Genetics:
A:Gene: GDB: TLE3; ESG; ESG3
A:Cross-references: GDB:228049; OMIM:600190
A:Map position: 15pter-15qter
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
C:Keywords: nucleus
F:482-514/Domain: WD repeat homology <WD1>
F:528-561/Domain: WD repeat homology <WD2>
F:614-647/Domain: WD repeat homology <WD3>
F:696-729/Domain: WD repeat homology <WD4>
F:734-770/Domain: WD repeat homology <WD5>

Query Match 19.2%; Score 58.5; DB 2; Length 772;
Best Local Similarity 38.3%; Pred. No. 71;

Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;

Qy 2 STKDFNLDLVSVSKKSGASPRITSLCTPGCKTGALMGCMKMTAT 48
Db 267 SPENGLDKARSLKKDAPTSVASVSSSTPSSKTDL-GHNDKST 312

RESULT 12

F84721
probable RING zinc finger protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84721
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Euse, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:1061197
A:Accession: F84721
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-565 <STO>
A:Cross-references: GB:A8002093; NID:94582446; PIDN:AAD24830.1; GSPDB:GND0139
C:Genetics:
A:Gene: At2g31510
A:Map position: 2

Query Match 19.1%; Score 58; DB 2; Length 565;
Best Local Similarity 28.8%; Pred. No. 61;

Matches 23; Conservative 9; Mismatches 20; Indels 28; Gaps 4;

Qy 3 TKDFN-----LDLVSVSK-----DSGASPRITSLCTPGC--KTGA 38
Db 426 SKDFNDFRTKLAGLSTVTKNPFENLVKALENGLADVDSDHAAASSTSTSGCCKSKRRE 485

Qy 39 LM-----GCMKMTATGCHSTH 54
Db 486 LVDPFGTKEPBLIVCNCTMH 505

RESULT 13

A56879
diacylglycerol kinase (EC 2.7.1.107) alpha - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 18-Jun-1999
C:Accession: A56879
R:Goto, K.; Watanabe, M.; Kondo, H.; Yuasa, H.; Sakane, F.; Kanoh, H.
Brain Res. Mol. Brain Res. 16, 75-87, 1992
A:Title: Gene cloning, sequence, expression and in situ localization of 80 kDa diacylglyc
A:Reference number: A56879; MUID:93095720; PMID:1339302
A:Contents: brain
A:Accession: A56879
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-727 <GOT>
A:Cross-references: GB:S49760; NID:9261423; PIDN:AAB24434.1; PID:9261424
C:Note: sequence extracted from NCBI backbone (NCBI:120083, NCBI:120084)
C:Superfamily: human diacylglycerol kinase; calmodulin repeat homology; protein kinase C
C:Keywords: ATP; calcium binding; duplication; EF hand; phosphotransferase; zinc
F:108-140/Domain: calmodulin repeat homology <EF1>
F:153-185/Domain: calmodulin repeat homology <EF2>
F:204-251/Domain: protein kinase C zinc-binding repeat homology <KZ1>
F:268-317/Domain: protein kinase C zinc-binding repeat homology <KZ2>

Query Match 18.9%; Score 57.5; DB 2; Length 727;
Best Local Similarity 24.6%; Pred. No. 87;

Matches 16; Conservative 6; Mismatches 28; Indels 15; Gaps 2;

Qy 1 MSTDFNLDLVSVSKKSGASPRITSLCTPGCKTGALMGCMKMTAT-----C 49
Db 246 MKAQCEVSTYAKSRKIDIGVPHVVRG-----GCHSGRCDBRCQKIRFYHSLTGHCVCW 301

RESULT 14

S35681
ESG protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-May-2000
C:Accession: S35681; S34162
R:Miyaasaka, H.; Choudhury, B.K.; Hou, E.W.; Li, S.S.L.
Eur. J. Biochem. 216, 343-352, 1993
A:Title: Molecular cloning and expression of mouse and human cDNA encoding AES and ESG pr
A:Reference number: S35678; MUID:93373944; PMID:8365415
A:Accession: S35681
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-771 <MTY>
A:Cross-references: EMBL:X73360; NID:931335; PIDN:CAAS1770.1; PID:9313336
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
C:Keywords: phosphoprotein
F:527-560/Domain: WD repeat homology <WD1>
F:613-646/Domain: WD repeat homology <WD3>
F:695-728/Domain: WD repeat homology <WD4>
F:736-769/Domain: WD repeat homology <WD5>

Query Match 18.9%; Score 57.5; DB 2; Length 771;
Best Local Similarity 38.3%; Pred. No. 92;

Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;

Qy 2 STKDFNLDLVSVSKKSGASPRITSLCTPGCKTGALMGCMKMTAT 48
Db 266 SPENGLDKARGLKKDAPTSVASVSSSTPSSKTDL-GHNDKST 311

RESULT 15

S70355
phosphoribosylanthranilate isomerase (EC 5.3.1.24) - Lipomyces starkeyi
C:Species: Lipomyces starkeyi
C:Date: 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S70355
R:Biggell, G.R.; Bruce, I.J.; Evans, I.H.
Curr. Genet. 30, 83-88, 1996
A>Title: Electrophoretic karyotype of the amylolytic yeast *Lipomyces starkeyi* and cloning
A:Reference number: S70355; MUID:96269934; PMID:8662214
A:Accession: S70355
A:Molecule type: DNA
A:Residues: 1-232 <BIG>
A:Cross-references: EMBL:Z68292; NID:G1134847; PIDN:CA92584.1; PID:G1161576
A>Note: the authors translated the codon GCC for residue 211 as Gln
C:Genetics:
A:Gene: TRP1
C:Superfamily: phosphoribosylanthranilate isomerase; trpF homology
C:Keywords: intramolecular oxidoreductase; isomerase; tryptophan biosynthesis
F:14-228/Domain: trpF homology <TRF>

Query Match 18.6%; Score 56.5; DB 2; Length 232;
Best Local Similarity 36.7%; Pred.No. 42;
Matches 11; Conservative 5; Mismatches 9; Indels 5; Gaps 1;

Ov 24 ITSTSLCTPGCKTGALMGCMKMTATCHCSI 53
Db 3 VSTSLCTPIVKI-----GGLTVEAARCAI 27

Search completed: January 12, 2004, 14:26:33
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2004, 14:18:07 ; Search time 18 Seconds
(without alignments)
148.918 Million cell updates/sec

Title: US-10-082-618-5
Perfect score: 304
Sequence: 1 MSTDFNLDLVSVSKDSCA.....ALMGCMKATCHCSIHVS 57

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	298	98.0	57	1 LANN_LACLA	P13068 lactococcus
2	291	95.7	57	1 LANN_LACLA	P23559 lactococcus
3	157	51.6	56	1 LANS_BACSU	P10946 bacillus su
4	94.5	31.1	52	1 LANS_STAEP	P08136 staphylococ
5	91.5	30.1	52	1 LANG_STAGA	P21838 staphylococ
6	72	23.7	46	1 SRTA_STRPY	P04642 streptococ
7	63	20.7	995	1 Y109_YEAST	P40442 saccharomyc
8	61	20.1	732	1 E4L5_HUMAN	P29315 rattus norv
9	59	19.4	456	1 RINI_RAT	P55947 helix pomat
10	58.5	19.2	64	1 MTCU_HELPO	P04726 homo sapien
11	58.5	19.2	772	1 TLE3_HUMAN	P51556 rattus norv
12	57.5	18.9	727	1 KODA_RAT	P09133 rattus norv
13	57.5	18.9	764	1 TLE3_MOUSE	P08122 mus musculu
14	57.5	18.9	771	1 TLE3_MOUSE	P97435 mus musculu
15	57.5	18.9	1069	1 ENTP_MOUSE	P01128 lipomyces s
16	56.5	18.6	232	1 TRPF_LIPST	P10775 sub scrofa
17	56.5	18.6	456	1 RINI_PIG	P06084 mus musculu
18	56.5	18.6	498	1 TNR8_MOUSE	P07141 rattus norv
19	56.5	18.6	741	1 TLE4_RAT	P06244 mus musculu
20	56.5	18.6	766	1 TLE4_MOUSE	P27591 drosophila
21	56.5	18.6	799	1 ITEN_DROME	P03376 chironomus
22	56.5	18.6	1700	1 BAR3_CHITE	P09166 homo sapien
23	56.5	18.6	3674	1 SPCK_HUMAN	P06586 streptococ
24	56	18.4	63	1 LANN_STRMU	P58913 cornus catus
25	56	18.4	71	1 CXOA_CONCT	P24043 homo sapien
26	56	18.4	3110	1 LMA2_HUMAN	P04727 homo sapien
27	55.5	18.3	766	1 TLE4_HUMAN	P01721 mus musculu
28	55	18.1	343	1 GAS1_MOUSE	P21551 amygdala m
29	55	18.1	369	1 WNT1_MOUSE	P23366 saccharomyc
30	55	18.1	374	1 ADH7_MOUSE	P09475 mus musculu
31	55	18.1	551	1 BEM1_YEAST	P09475 mus musculu
32	55	18.1	715	1 S21F_MOUSE	P09475 mus musculu
33	55	18.1	716	1 S21F_RAT	P09475 mus musculu

34	54.5	17.9	64	1 MT_STENE	P55953 sterechilus
35	54.5	17.9	2193	1 POLG_HE71B	P06478 h genome po
36	54	17.8	166	1 TH12_RAT	P97615 rattus norv
37	54	17.8	511	1 KPYC-SOVEN	P42805 glycine max
38	54	17.8	808	1 MCM3_HUMAN	P25205 homo sapien
39	54	17.8	876	1 TOP1_VIBCH	P09162 vibrio chol
40	54	17.8	1257	1 RBBI_HUMAN	P29374 homo sapien
41	54	17.8	2193	1 POLG_CX16G	P05900 c genome po
42	53.5	17.6	166	1 TH12_HUMAN	P09757 homo sapien
43	53.5	17.6	442	1 TIG_CHLTR	P04713 chlamydia t
44	53.5	17.6	3672	1 LML2_CAERL	P01313 caenorhabdi
45	53	17.4	298	1 Y1E4_YEAST	P40529 saccharomyc

ALIGNMENTS

RESULT 1	ID	LANN_LACLA	STANDARD;	PRT;	57 AA.
AC	P13068:	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Lanthibiotic nisin A precursor.				
GN	SPAN OR NISA.				
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.				
OX	NCBI_TaxID=1360;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 11454 / DSM 20729 / NCDO 496;				
RX	MEDLINE=88034093; PubMed=3141403;				
RA	Butcher G.W., Banerjee S., Hansen J.N.;				
RT	"Structure, expression, and evolution of a gene encoding the				
RT	precursor of nisin, a small protein antibiotic.";				
RL	J. Biol. Chem. 263:16260-16266 (1988).				
RN	(2)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 11454 / DSM 20729 / NCDO 496;				
RX	MEDLINE=91282469; PubMed=1905517;				
RA	Steen M.T., Chung Y.J., Hansen J.N.;				
RT	"Characterization of the nisin gene as part of a polycistronic operon				
RT	in the chromosome of Lactococcus lactis ATCC 11454.";				
RL	Appl. Environ. Microbiol. 57:1181-1188 (1991).				
RN	(3)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=6F3;				
RX	MEDLINE=89155467; PubMed=2493449;				
RA	Kaletka C., Entian K.-D.;				
RT	"Nisin, a peptide antibiotic: cloning and sequencing of the nisa gene				
RT	and posttranslational processing of its peptide product.";				
RL	J. Bacteriol. 171:1597-1601 (1989).				
RN	(4)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=6F3;				
RX	MEDLINE=93128945; PubMed=1482192;				
RA	Engelke G., Gutowski-Eckel Z., Hammelmann M., Entian K.-D.;				
RT	"Biosynthesis of the lantibiotic nisin: genomic organization and				
RT	membrane localization of the NisB protein.";				
RL	Appl. Environ. Microbiol. 58:3730-3743 (1992).				
RN	(5)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=N120 R5;				
RX	MEDLINE=93373937; PubMed=7689965;				
RA	Kuipers O.P., Beertuyzen M.M., Sizeren R.J., de Vos W.M.;				
RT	"Characterization of the nisin gene cluster nisaB-CIR of Lactococcus				
RT	lactis. Requirement of expression of the nisa and nisa genes for				
RL	development of immunity.";				
RN	Eur. J. Biochem. 216:281-291 (1993).				
RP	(6)				
RC	SEQUENCE OF 24-57.				
RA	Gross E.;				

RL (in) Friedman M. (eds.);
 RL Protein cross-linking, pp.131-153, Plenum Press, New York (1977).
 RN [7]
 RP SEQUENCE OF 24-57.
 RX MEDLINE=72072901; PubMed=5131162;
 RA Gross E., Morell J.L.;
 RT "The structure of nisin";
 RN J. Am. Chem. Soc. 93:4634-4635(1971).
 [8]
 RP STRUCTURE BY NMR.
 RX MEDLINE=92111494; PubMed=1765078;
 RA van de Ven F.J., van den Hooven H.W., Konings R.N.H., Hilbers C.W.;
 RT "NMR studies of lantibiotics. The structure of nisin in aqueous
 solution";
 RN Eur. J. Biochem. 202:1181-1188(1991).
 [9]
 RP STRUCTURE BY NMR.
 RX MEDLINE=92246867; PubMed=1575686;
 RA Llan L.-Y., Chan W.C., Morley S.D., Roberts G.C.K., Bycroft B.W.,
 RA Jackson D.;
 RT "Solution structures of nisin A and its two major degradation
 products determined by NMR";
 RN Biochem. J. 283:413-420(1992).
 [10]
 RP STRUCTURE BY NMR.
 RX MEDLINE=93202265; PubMed=8454055;
 RA van den Hooven H.W., Fogolari F., Rollemans H.S., Konings R.N.H.,
 RA Hilbers C.W., van de Ven F.J.;
 RT "NMR and circular dichroism studies of the lantibiotic nisin in non-
 aqueous environments";
 RN FEBS Lett. 319:189-194(1993).
 [11]
 RP STRUCTURE BY NMR.
 RX MEDLINE=93120109; PubMed=8418850;
 RA Sailer W., Helms G.L., Henkel T., Niemczura W.P., Stiles M.E.,
 RA Vederas J.C.;
 RT "15N- and 13C-labeled media from *Anabaena* sp. for universal isotopic
 labeling of bacteriologins: NMR resonance assignments of leucococcin A
 from *Leuconostoc gelidium* and nisin A from *Lactococcus lactis*";
 RN Biochemistry 32:310-318(1993).
 [12]
 RP FUNCTION: LANTHONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
 CC ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
 CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
 CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
 CC TRANSMEMBRANE PORES.
 CC -1- PTM: Maturation of lantibiotics involves the enzymic conversion of
 CC Thr, and Ser into dehydrated AA and the formation of thioether
 CC bonds with cysteine. This is followed by membrane translocation
 CC and cleavage of the modified precursor.
 CC -1- MISCELLANEOUS: USED AS A FOOD PRESERVATIVE.
 CC PLASMID DNA. THE SEQUENCES REPORTED ARE ABSOLUTELY IDENTICAL.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL substation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sib.ch).
 CC -----
 DR EMBL, J04057; AAA88606.1; -
 DR EMBL, M65089; AAA73038.1; -
 DR EMBL, M24527; AAA26948.1; -
 DR EMBL, X68307; CAA48380.1; -
 DR EMBL, M27277; AAA25188.1; -
 DR EMBL, D00696; BAA00602.1; -
 DR EMBL, L16226; AAA25189.1; -
 DR EMBL, M79445; AAA25198.1; -
 DR PIR, A31915; NILSA.
 DR InterPro; IPR006079; Lan_dcm.
 DR InterPro; IPR000446; Nisin.

DR Pfam; PF02052; Gallidermin; 1.
 DR PRINTS; PR00324; NISIN.
 KM Antibiotic; Bacteriocin; Lantibiotic; D-amino acid; Plasmid;
 KM Thioether bond.
 FT PROPEP 1 23
 FT CHAIN 24 57 LANTIBIOTIC NISIN A.
 FT CROSSLINK 26 30 Lanthionine (Ser-Cys).
 FT CROSSLINK 31 34 Beta-methylanthionine (Thr-Cys).
 FT CROSSLINK 36 42 Beta-methylanthionine (Thr-Cys).
 FT CROSSLINK 46 49 Beta-methylanthionine (Thr-Cys).
 FT CROSSLINK 48 51 Beta-methylanthionine (Thr-Cys).
 FT MOD RES 25 25 DHB (2,3-DIDEHYDROBUTYRINE).
 FT MOD RES 28 28 DHA (2,3-DIDEHYDROALANINE).
 FT MOD RES 56 56 DHA (2,3-DIDEHYDROALANINE).
 SQ SEQUENCE 57 AA; 5963 MW; 315E428AC70BF8A CRC64;
 Query Match 98.0%; Score 298; DB 1; Length 57;
 Best Local Similarity 98.2%; Pred. No. 8.2e-31;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSTDNFNDLVSVSKDGSASPRITSTLCPTGCKTGALMGCMKMTATCHCSIHYSK 57
 DB 1 MSTDNFNDLVSVSKDGSASPRITSTLCPTGCKTGALMGCMKMTATCHCSIHVS 57
 RESULT 2
 LANTZ_LACIA STANDARD; PRT; 57 AA.
 ID LANTZ_LACIA
 AC P29559;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lantibiotic nisin Z precursor.
 GN NISZ.
 OS *Lactococcus lactis* (subsp. *lactis*) (*Streptococcus lactis*).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; *Lactococcus*.
 OX NCBI_Taxid=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIZO 22186;
 RX MEDLINE=92037612; PubMed=1935953;
 RA Mulders J.W.M., Boerrigter I.J., Rollemans H.S., Sieszen R.J.,
 RA de Vos W.M.;
 RT "Identification and characterization of the lantibiotic nisin Z, a
 RT natural nisin variant";
 RN Eur. J. Biochem. 201:581-584(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 7638;
 RX Araya T., Ishibashi N., Shimamura S.;
 RA "Genetic evidence that *Lactococcus lactis* JCM7638 produces a mutated
 RT form of nisin";
 RN J. Gen. Appl. Microbiol. 38:271-278(1992).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N8;
 RX MEDLINE=95352820; PubMed=7626780;
 RA Immonen T., Ye S., Ra R., Qiao M., Paulin L., Saxis P.E.J.;
 RT "The codon usage of the nisz operon in *Lactococcus lactis* N8 suggests
 RT a non-lactococcal origin of the conjugative nisin-sucrose
 RT transposon";
 RN DNA Seq. 5:203-218(1995).
 [4]
 RP FUNCTION: LANTHONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
 CC ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
 CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
 CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
 CC TRANSMEMBRANE PORES.
 CC -1- PTM: Maturation of lantibiotics involves the enzymic conversion of
 CC Thr, and Ser into dehydrated AA and the formation of thioether
 CC bonds with cysteine. This is followed by membrane translocation
 CC and cleavage of the modified precursor.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X61144; CAA43440.1; -
 DR EMBL; D10768; BAA01598.1; -
 DR EMBL; Z18947; CAA79467.1; -
 DR InterPro; IPR006079; Lan_dom.
 DR Pfam; PF02052; Gallidermin; 1.
 DR PRINTS; PR00324; Nisin.
 DR Antibiotic; Bacteriocin; Lantibiotic; D-amino acid; Thioether bond.
 KM PROPEP 1 23
 FT CHAIN 24 57
 FT CROSSLINK 26 30 LANTIBIOTIC NISIN Z.
 FT CROSSLINK 31 34 Lanthionine (Ser-Cys).
 FT CROSSLINK 36 42 Beta-methylanthionine (Thr-Cys).
 FT CROSSLINK 46 49 Beta-methylanthionine (Thr-Cys).
 FT CROSSLINK 48 51 Beta-methylanthionine (Thr-Cys).
 FT MOD_RES 25 25 DHB (2,3-DIDEHYDROBUTYRINE).
 FT MOD_RES 28 28 DHA (2,3-DIDEHYDROALANINE).
 FT MOD_RES 56 56 N -> H (IN STRAIN JCM7638).
 FT VARIANT 50 50 N -> H (IN STRAIN JCM7638).
 SQ SEQUENCE 57 AA; 5940 MW; D5E84428ACT0BFEE CRC64;

Query Match 95.7%; Score 291; DB 1; Length 57;
 Best Local Similarity 96.5%; Pred. No. 6.2e-30;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTKDFNLDVSVSKKDSGASPRITSTICTPGCKTGALMGCMKATCNCISHSK 57
 DB 1 MSTKDFNLDVSVSKKDSGASPRITSTICTPGCKTGALMGCMKATCNCISHSK 57

RESULT 3
 LANS_BACSU STANDARD; PRT; 56 AA.
 AC P10946;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lantibiotic subtilin precursor.
 GN SPAS OR SUB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid:1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88243844; PubMed=2837490;
 RA Banerjee S., Hansen J.N.;
 RT "Structure and expression of a gene encoding the precursor of
 RT subtilin, a small protein antibiotic.";
 RL J. Biol. Chem. 263:9508-9514 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 6633 / LH45;
 RX MEDLINE=92138640; PubMed=1735728;
 RA Chung Y.J., Steen M.T., Hansen J.N.;
 RT "The subtilin gene of Bacillus subtilis ATCC 6633 is encoded in an
 RT operon that contains a homolog of the hemolysin B transport
 RT protein.";
 RL J. Bacteriol. 174:1417-1422 (1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 6633 / LH45;
 RX MEDLINE=92171481; PubMed=153969;
 RA Klein C., Kaletta C., Schnell N., Entian K.-D.;
 RT "Analysis of genes involved in biosynthesis of the lantibiotic
 RT subtilin.";

RL Appl. Environ. Microbiol. 58:132-142 (1992).
 RN [4]
 RP SEQUENCE OF 25-56.
 RX MEDLINE=75040028; PubMed=4154277;
 RA Gross E., Kiltz H.H., Nebelin E.;
 RT "Subtilin, VI: the structure of subtilin.";
 RL Hope-Seyler's Z. Physiol. Chem. 354:810-812 (1973).
 RN [5]
 RP MODE OF ACTION.
 RX MEDLINE=89276381; PubMed=2471644;
 RA Scheller F., Benz R., Sahl H.-G.;
 RT "The peptide antibiotic subtilin acts by formation of
 RT voltage-dependent multi-state pores in bacterial and artificial
 RT membranes.";
 RL Eur. J. Biochem. 182:181-186 (1989).
 RN [6]
 RP STRUCTURE BY NMR.
 RC STRAIN=ATCC 6633 / LH45;
 RX MEDLINE=92192284; PubMed=1547888;
 RA Chan W.C., Bycroft B.W., Leylands M.L., Lian L.-Y., Yang J.C.,
 RA Roberts G.C.K.;
 RT "Sequence-specific resonance assignment and conformational analysis
 RT of subtilin by 2D NMR.";
 RL FEBS Lett. 300:56-62 (1992).
 RN [7]
 RP MUTAGENESIS OF SER-29.
 RX MEDLINE=93167833; PubMed=8434932;
 RA Liu W., Hansen J.N.;
 RT "The antimicrobial effect of a structural variant of subtilin against
 RT outgrowing Bacillus cereus T spores and vegetative cells occurs by
 RT different mechanisms.";
 RL Appl. Environ. Microbiol. 59:648-651 (1993).
 CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
 CC ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
 CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
 CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
 CC TRANSMEMBRANE PORES.
 CC -1- PTM: Maturation of lantibiotics involves the enzymic conversion of
 CC Thr, and Ser into dehydrated AA and the formation of thioether
 CC bonds with cysteine. This is followed by membrane translocation
 CC and cleavage of the modified precursor.
 CC -1- MISCELLANEOUS: SUBTILIN ACTIVITY IS OBSERVED DURING STATIONARY
 CC PHASE, BUT NOT DURING EXPONENTIAL GROWTH.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER LANTIBIOTICS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; J03767; AAA22841.1; -
 DR EMBL; M86869; AAA22840.1; -
 DR EMBL; M83944; AAA22772.1; -
 DR EMBL; M92633; AAA22778.1; -
 DR EMBL; U09619; AAB91589.1; -
 DR PIR; A28112; NIBSSA.
 DR InterPro; IPR006079; Lan_dom.
 DR InterPro; IPR000446; Nisin.
 DR Pfam; PF02052; Gallidermin; 1.
 DR PRINTS; PR00324; Nisin.
 DR Antibiotic; Bacteriocin; Lantibiotic; D-amino acid; Thioether bond.
 KM PROPEP 1 24
 FT CHAIN 25 56
 FT CROSSLINK 27 31 LANTIBIOTIC SUBTILIN.
 FT CROSSLINK 32 35 Lanthionine (Ser-Cys).
 FT CROSSLINK 37 43 Beta-methylanthionine (Thr-Cys).
 FT CROSSLINK 47 50 Beta-methylanthionine (Thr-Cys).
 FT CROSSLINK 49 52 Beta-methylanthionine (Thr-Cys).
 FT MOD_RES 29 29 DHA (2,3-DIDEHYDROALANINE).
 FT MOD_RES 42 42 DHB (2,3-DIDEHYDROBUTYRINE).
 FT MOD_RES 42 42

FT MOD_RES 55 55 DHA (2,3-DIDEHYDROALANINE).
 FT MUTAGEN 29 29 S->A: DEVOID OF ANTIMICROBIAL ACTIVITY;
 FT SEQUENCE 56 AA; 6218 MW; DA9707FBE8A1EBBA CRC64;
 SQ SEQUENCE 56 AA; 6218 MW; DA9707FBE8A1EBBA CRC64;
 Query Match 51.6%; Score 157; DB 1; Length 56;
 Best Local Similarity 59.2%; Pred. No. 3, 6e-13;
 Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;
 QY 5 DNLDLVSVSKRSGASPRITSTSLCTPGCKTGALMCMKATGCSI 53
 DB 6 DFDLDVVKVSKODSKITPQWKSBSLCTPGCVGATQTCPLQTLTCNCKI 54
 RESULT 4
 LANE_STAMP STANDARD; PRT; 52 AA.
 ID LANE_STAMP STANDARD; PRT; 52 AA.
 AC P08136; Q54093;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lanthibiotic epidermin precursor.
 GN EPIA.
 OS Staphylococcus epidermidis.
 OG Plasmid pTU 32.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_Taxid=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TU 3298 / DSM 3095;
 RX MEDLINE=88216821; PubMed=2835685;
 RA Schnell N., Entian K.-D., Schneider U., Gotz F., Zahner H.,
 RA Kellner R., Jung G.;
 RT "Prepeptide sequence of epidermin, a ribosomally synthesized
 RT antibiotic with four sulphide-rings."
 RT Nature 333:276-278 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TU 3298 / DSM 3095;
 RX MEDLINE=92155237; PubMed=1740156;
 RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
 RA Goetz F., Entian K.-D.;
 RT "Analysis of genes involved in the biosynthesis of lantibiotic
 RT epidermin."
 RT Eur. J. Biochem. 204:57-68 (1992).
 RN [3]
 RP FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
 RP ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
 RP LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
 RP CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
 RP TRANSMEMBRANE PORES.
 RP -1- PTM: Maturation of lantibiotics involves the enzymic conversion of
 RP Thr. and Ser into dehydrated AA and the formation of thioether
 RP bonds with cysteine. The C-terminal lantionine undergoes
 RP decarboxylation. This is followed by membrane translocation and
 RP cleavage of the modified precursor.
 RP -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
 RP This SWISS-PROT entry is copyright. It is produced through a collaboration
 RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
 RP the European Bioinformatics Institute. There are no restrictions on its
 RP use by non-profit institutions as long as its content is in no way
 RP modified and this statement is not removed. Usage by and for commercial
 RP entities requires a license agreement (See http://www.isb-sib.ch/announce/
 RP or send an email to license@isb-sib.ch).
 CC EMBL; X07840; CAA30689.1; -
 CC EMBL; X07840; CAA30690.1; -
 CC EMBL; X62386; CAA44252.1; -
 CC EMBL; A12927; CAA01070.1; -
 CC PIR; S00768; EISED.
 CC PDB; 1G5Q; 02-MAY-01.
 CC InterPro; IPR006078; Gallidermin.
 CC InterPro; IPR006079; lan_dom.

DR Pfam; PF02052; Gallidermin; 1.
 DR PRINTS; PR00323; GALLIDERMIN.
 KW Antibiocic; Bacteriocin; Lantibiotic; D-amino acid; Plasmid;
 KW 3D-structure; Thioether bond.
 FT PROPEP 1 30
 FT CHAIN 31 52 LANTIBIOTIC EPIDERMIN.
 FT CROSSLINK 33 37 Lanthionine (Ser-Cys).
 FT CROSSLINK 38 41 Beta-methylanthionine (Thr-Cys).
 FT CROSSLINK 46 51 Lanthionine (Ser-Cys).
 FT CROSSLINK 49 52 S-(2-aminovinyl)-D-cysteine (Ser-Cys).
 FT MOD_RES 44 44 DHB (2,3-DIDEHYDROBUTYRINE).
 SQ SEQUENCE 52 AA; 5632 MW; 8B1AD2875BF16D6D CRC64;
 Query Match 31.1%; Score 94.5; DB 1; Length 52;
 Best Local Similarity 62.2%; Pred. No. 2, 2e-05;
 Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;
 QY 6 FNLDLVSVSKR--DSGASPRITSTSLCTPGC-KTGA 38
 DB 11 FNLD-VKVNAKESNDGAEPRITASKICTPGCAKTGS 46
 RESULT 5
 LANG_STAGA STANDARD; PRT; 52 AA.
 ID LANG_STAGA STANDARD; PRT; 52 AA.
 AC P21838;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lanthibiotic gallidermin precursor.
 GN GDMA.
 OS Staphylococcus gallinarum.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_Taxid=1293;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=89306540; PubMed=2765032;
 RA Schnell N., Entian K.-D., Goetz F., Hoerner T., Kellner R., Jung G.;
 RT "Structural gene isolation and prepeptide sequence of gallidermin, a
 RT new lantionine containing antibiotic."
 RT FEWS Microbiol. Lett. 49:263-267 (1989).
 RN [2]
 RP SEQUENCE OF 31-52.
 RC STRAIN=TU 3928;
 RX MEDLINE=89030695; PubMed=3181159;
 RA Kellner R., Jung G., Hoerner T., Zaehner H., Schnell N., Entian K.-D.,
 RA Goetz F.;
 RT "Gallidermin: a new lantionine-containing polypeptide antibiotic."
 RT Eur. J. Biochem. 177:53-59 (1988).
 RN [3]
 RP STRUCTURE BY NMR.
 RC MEDLINE=92032577; PubMed=1932575;
 RA Freund S., Jung G., Gutbrod O., Folkers G., Gibbons W.A., Allgater H.,
 RA Werner R.;
 RT "The solution structure of the lantibiotic gallidermin."
 RT Biopolymers 31:803-811 (1991).
 RN [4]
 RP FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
 RP ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF
 RP LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
 RP CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
 RP TRANSMEMBRANE PORES.
 RP -1- PTM: Maturation of lantibiotics involves the enzymic conversion of
 RP Thr. and Ser into dehydrated AA and the formation of thioether
 RP bonds with cysteine. The C-terminal lantionine undergoes
 RP decarboxylation. This is followed by membrane translocation and
 RP cleavage of the modified precursor.
 RP -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
 RP PIR; A61072; EPSGD.
 RP InterPro; IPR006078; Gallidermin.
 RP InterPro; IPR006079; lan_dom.
 RP Pfam; PF02052; Gallidermin; 1.
 DR PRINTS; PR00323; GALLIDERMIN.
 KW Antibiotic; Bacteriocin; Lantibiotic; D-amino acid; Thioether bond.

```

FT PROPER 1 30 LANTIBIOTIC GALLDERMIN.
FT CHAIN 31 52
FT CROSSLINK 33 37 Lanthionine (Ser-Cys).
FT CROSSLINK 38 41 Beta-methylanthionine (Thr-Cys).
FT CROSSLINK 46 51 Lanthionine (Ser-Cys).
FT CROSSLINK 49 52 S-(2-aminoviny1)-D-cysteine (Ser-Cys).
FT MOD RES 44 44 DHB (2,3-DIDEHYDROBUTYRINE).
SQ SEQUENCE 52 AA; 5647 MW; 8584C0040AB4786D CRC64;

Query Match 30.1%; Score 91.5; DB 1; Length 52;
Best Local Similarity 62.2%; Pred. No. 5,2e-05;
Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;

Db 11 FDL-D-VKYNAKESNDGAEPRIASKFLCTPGCAKTGS 46

RESULT 6
SRTA_STRPY STANDARD; PRT; 46 AA.
ID SRTA_STRPY
AC Q9PDV1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lantibiotic srta precursor.
GN SRTA OR SPY1083.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;

[1]
SEQUENCE FROM N.A.
RC STRAIN=BL-T;
RA Karaya K., Takeo A.;
RT "Gene cluster of lantibiotics producing by Streptococcus pyogenes.";
RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.

[2]
SEQUENCE FROM N.A.
RP STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; Pubmed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezele S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RL "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB030831; BAB08162.1; -
DR EMBL; AE006552; AAK33966.1; -
KW Antibiotic; Bacteriocin; Lantibiotic; Complete proteome.
FT PROPEP 1 ? POTENTIAL.
FT CHAIN ? 46 LANTIBIOTIC SRTA.
SQ SEQUENCE 46 AA; 5219 MW; 3775CC54B8A2B00F CRC64;

Query March 23.7%; Score 72; DB 1; Length 46;
Best Local Similarity 56.7%; Pred. No. 0.013;
Matches 17; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

Db 4 KDFNLDLVSVKSGASPRITSTSLCPG 33
6 KDFDLDL-KTNKQDT-ATPYGSRLLCTPG 33

```

ID	Y109_YEAST	STANDARD	PRT	995 AA.
AC	P40442.1			
DT	01-FEB-1995	(Rel. 31, Created)		
DT	01-FEB-1995	(Rel. 31, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	Hypothetical 99.7 kDa protein in Sdu1 5' region precursor.			
CN	YIL169C OR Y19402.07C			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_Taxid=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SD288C / AB972;			
RX	PubMed=9169870;			
RA	Churcher C.M., Bowman S., Baddock K., Bankier A., Brown D.,			
RA	Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,			
RA	Harris D.E., Horenell T., Hunt S., Jagsals K., Jones M., Lye G.,			
RA	Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,			
RA	Skeilton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;			
RT	"The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";			
RL	Nature 387:84-87(1997).			
CC	-1- SIMILARITY: Contains 1 methyl-accepting transducer domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.ebi.ac.uk/submit/			
CC	or send an email to license@ebi.ac.uk).			
DR	EMBL; Z46921, CAAB7023.1; -.			
DR	PIR; S50358; S50358.			
DR	SGD; S0001431; YIL169C.			
DR	InterPro; IPR004089; Clntaxis.transd.			
DR	PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2_1.			
KW	Hypothetical protein; Signal.			
FT	SIGNAL	1	23	POTENTIAL.
FT	CHAIN	24	995	HYPOTHETICAL PROTEIN YIL169C.
FT	DOMAIN	26	253	METHYL-ACCEPTING TRANSDUCER.
FT	CARBOHYD	28	28	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	35	35	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	468	468	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	664	664	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	995 AA;	99735 MW;	FE3E287A03F137EC CRC64;
Query Match		20.7%;	Score 63;	DB 1; Length 995;
Best Local Similarity		26.9%;	Pred. No. 4.6;	
Matches	21;	Conservative	8;	Mismatches 25; Indels 24; Gaps 2;
Oy	3	TKPDNDLVSVSKDSDASPRITSTSLCTGCGC-----ITGGLM-----		40
Db	676	TTDSNGENVYITTTTPCSTTATITSCDETGCHVSTSGAVETVSSKSYTATVTHCD		735
Oy	41	--GCNMKTATCHCSIHVS	56	
Db	736	DNGCNTVITVSECKETS	753	
RESULT 8				
BALS_HUMAN				
ID	EAL5_HUMAN	STANDARD;	PRT;	732 AA.
AC	Q9HGM4; Q9H975;			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	Band 4.1-like protein 5.			
CN	EPB41L5 OR KIA1548.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahashi K., Masuno Y., Sasaki N.;
RT "NEO human cDNA sequencing project.";
RN Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE OF 188-732 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Ref. 71273-281(2000).
CC -1- SIMILARITY: Contains 1 FERM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AK03019; BAB14360.1; -
DR EMBL; AB046768; BAB13374.1; -
DR Gene; HGNC:19819; EBP415.
DR InterPro; IPR000299; Band_4.1.
DR Pfam; PF00373; Band_4; 1.
DR PRINTS; PR00935; BAND41.
DR SMART; SM00295; B41; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS50057; FERM_3; 1.
DR Cytokeleton.
KW DOMAIN
FT 43 337 FERM.
FT 669 732 SCAMSLAGCEMLTGKRGHGKNGKDISLSPAPPLUDAV
FT TSSGPIAEAVTKKCKLITTEL -> LMSHGRRSCPEAE
FT VFTDH (IN REF. 1).
SQ SEQUENCE 732 AA; 81758 MW; 76D5BD8CE09E761 CRC64;

Query Match 20.1%; Score 61; DB 1; Length 732;
Best Local Similarity 29.4%; Pred. No. 5.9;
Matches 15; Conservative 12; Mismatches 16; Indels 8; Gaps 2;

QY 2 STRDPLDLVSVSK--DSGASPRITSLCTP-----GCKTGALMGCM 44
DB 631 ATDELDALSLTENLIDHTVAQVSTSMITPRWVPGAMSNGLAGCM 681

RESULT 9
RINI_RAT STANDARD; PRT; 456 AA.
AC P29315;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease inhibitor.
GN RNH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=92162755; PubMed=1356887;
RA Kawamoto M., Mochizuma K., Sasaki M., Hattori H., Goto S.;
RT "cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue

```

```

RT distribution of the mRNA.";
RL Blochim. Biophys. Acta 1129:335-338(1992).
CC -1- FUNCTION: INHIBITOR OF PANCREATIC RNASE AND ANGIOGENIN. MAY ALSO
CC -1- FUNCTION IN THE MODULATION OF CELLULAR ACTIVITIES.
CC -1- SUBUNIT: FORMS A TIGHT ONE-TO-ONE COMPLEX WITH THE RNASE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: BRAIN, HEART, LUNG, LIVER, SPLEEN, TESTES AND
CC KIDNEY. HIGHEST IN THE LUNG AND LOWEST IN THE HEART.
CC -1- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X62528; CAA44388.1; -
DR PIR; S20597; S20597.
DR HSSP; P10775; 2BNH.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR Pfam; PF00560; LRR; 4.
DR SMART; SM00368; LRR_R1; 1.
KW Repeat; Leucine-rich repeat.
FT REPEAT 15 43
FT REPEAT 44 71 LRR A1.
FT REPEAT 72 100 LRR A2.
FT REPEAT 101 128 LRR B2.
FT REPEAT 129 157 LRR A3.
FT REPEAT 158 185 LRR B3.
FT REPEAT 186 214 LRR A4.
FT REPEAT 215 242 LRR B4.
FT REPEAT 243 271 LRR A5.
FT REPEAT 272 299 LRR B5.
FT REPEAT 300 328 LRR A6.
FT REPEAT 329 356 LRR B6.
FT REPEAT 357 385 LRR A7.
FT REPEAT 386 413 LRR B7.
FT REPEAT 414 442 LRR A8.
SQ SEQUENCE 456 AA; 49905 MW; 8518E5B1F09E598 CRC64;

Query Match 19.4%; Score 59; DB 1; Length 456;
Best Local Similarity 34.0%; Pred. No. 6.3;
Matches 17; Conservative 8; Mismatches 19; Indels 6; Gaps 3;

QY 8 LDVSVSKKDSGASPRITSLCTPGCKTGAL--MGCMKRTATC--HCSI 53
DB 284 LSLAGNELKDEGA--QLLCESLIERGCOLLESILMWKTSLTAA SCPHFSV 331

RESULT 10
MTCU_HELPO STANDARD; PRT; 64 AA.
AC P55947;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Copper-metallothionein (Cu-MT).
OS Helix pomatia (Roman snail) (Edible snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigamurethra; Helicoidea; Helicidae; Helix.
OX NCBI_TaxID=6536;
RN [1]
RP SEQUENCE.
RC TISSUE=Mantle;
RX MEDLINE=97373947; PubMed=9230430;
RA Dallinger R., Berger B., Hunziker P.E., Kaegi J.H.R.;
RT "Metallothionein in snail Cd and Cu metabolism.";
RL Nature 388:237-238(1997).
CC -1- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR

```

SEQUESTRATION OF TOXIC METAL IONS AND REGULATION OF ESSENTIAL TRACE ELEMENTS. THIS ISOFORM BINDS EXCLUSIVELY COPPER.

-1- DOMAIN 14' CYSTEINE RESIDUES ARE ARRANGED IN C-X-C GROUPS. THESE ARE THOUGHT TO BE THE METAL-BINDING SITES IN OTHER METALLOTHIONEINS.

-1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.

HSSP; P05106; 1JY2.

DR InterPro: IPR002400; GF cystknot.

DR PRINTS: PRO0438; GFCYSKNOT.

KW Metal-binding; Metal-thiolate cluster; Copper; Acetylation.

FT MOD_RES 1 1 ACETYLATION.

FT METAL 7 7 COPPER.

FT METAL 11 11 COPPER.

FT METAL 16 16 COPPER.

FT METAL 18 18 COPPER.

FT METAL 22 22 COPPER.

FT METAL 24 24 COPPER.

FT METAL 28 28 COPPER.

FT METAL 30 30 COPPER.

FT METAL 33 33 COPPER.

FT METAL 36 36 COPPER.

FT METAL 38 38 COPPER.

FT METAL 43 43 COPPER.

FT METAL 45 45 COPPER.

FT METAL 49 49 COPPER.

FT METAL 55 55 COPPER.

FT METAL 57 57 COPPER.

FT METAL 61 61 COPPER.

FT METAL 63 63 COPPER.

SO SEQUENCE 64 AA; 6205 MW; 96CC1998B7E12297 CRC64;

Query Match 19.2%; Score 58.5; DB 1; Length 64;

Best Local Similarity 50.0%; Pred. No. 0.87;

Matches 12; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

Qy 30 CTPGCKTALMGCMN-KTATCHCS 52

Db 18 GNDCKCGA--GNCNDRSSCHCS 39

RESULT 11

TL33 HUMAN STANDARD; PRT; 772 AA.

AC 004726; O8IVV6; O8WVR2; O9HGM5;

DT 01-FEB-1994 (Rel. 28, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Transducin-like enhancer protein 3 (ESG3).

GN TL33 OR KIAA1547.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCB1_TaxID=9606;

OX NCB1_TaxID=9606;

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Fetal brain;

RX MEDLINE=93265135; PubMed=1303260;

RA Stifani S., Blaumel C.M., Redhead N.J., Hill R.E.,

RA Aravanis-Tsakonas S.;

RT "Human homologs of a Drosophila enhancer of split gene product define a novel family of nuclear proteins."

RL Nat. Genet. 21:119-127(1992).

RM [2]

RN SEQUENCE FROM N.A. (ISOFORM 2).

RP TISSUE=Brain;

RC MEDLINE=20450683; PubMed=1097877;

RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XVIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."

RL DNA Res. 7:273-281(2000).

RM [3]

RN SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).

RC TISSUE=Pancreas, and Testis;

RX MEDLINE=22388257; PubMed=12477932;

RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,

RA Altshul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bat N.K.,

RA Hopkin R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Caavaant T.L., Scheetz T.E.,

RA Brownstein M.J., Uadin T.B., Toshiyuki S., Cantoni P., Prange C.,

RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Boeak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RM [4]

RN SEQUENCE OF 324-338 AND 521-531, AND INTERACTION WITH FOXA2.

RP MEDLINE=20309797; PubMed=10748198;

RA Wang J.-C., Walther-Law M., Yamada K., Osawa H., Stifani S.,

RA Granner D.K.,

RT "Transducin-like enhancer of split proteins, the human homologs of Drosophila groucho, interact with hepatic nuclear factor 3beta.";

RL J. Biol. Chem. 275:18418-18423(2000).

CC -1- FUNCTION: Transcriptional corepressor that binds to a number of transcription factors. Inhibits the transcriptional activation mediated by CTNBL1 and TCF family members in Wnt signaling. The effects of full-length TLE family members may be modulated by association with dominant-negative AES (By similarity).

CC -1- SUBUNIT: Homooligomer and heterooligomer with other family members. Binds LEF1, TCF7L1 and TCF7L2 (By similarity).

CC Binds FOXA2.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=4;

CC Comment=Experimental confirmation may be lacking for some isoforms;

CC Name=1;

CC IsoId=Q04726-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q04726-2; Sequence=VSP_006788;

CC Name=3;

CC IsoId=Q04726-3; Sequence=VSP_006789, VSP_006790;

CC Name=4;

CC IsoId=Q04726-4; Sequence=VSP_007023, VSP_007024, VSP_006790;

CC -1- TISSUE SPECIFICITY: Placenta and lung.

CC -1- SIMILARITY: Contains 7 WD repeats.

CC -1- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; M99438; AAA61194.1; -

CC DR EMBL; AB046767; BAB13373.1; ALT INIT.

CC DR EMBL; BC015729; -; NOT ANNOTATED_CDS.

CC DR EMBL; BC041831; AAA41831.1; -

CC DR PIR; D56695; D56695.

CC DR GeneW; HGNC:11839; TLE3.

CC DR MIM; 600190; -

CC DR GO; GO:0005634; C:nucleus; TNS.

CC DR GO; GO:0007397; P:histogenesis and organogenesis; TNS.

Db 302 HLEIH.306

RESULT 13
TLE3_RAT STANDARD; PRT; 764 AA.
AC 09JIT3;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transducin-like enhancer protein 3 (rtle3).
GN TLE3 OR ESP3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OK NCBI_TaxId=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex;
RX MEDLINE=20258864; PubMed=1080926;
RA Feldman J.D., Vician L., Crispino M., Hoe W., Baudry M.,
RA Herschman H.R.;
RT "rtle3, a newly identified transducin-like enhancer of split, is
RT induced by depolarization in brain."
RL J. Neurochem. 74:1838-1847(2000).
CC -I- FUNCTION: Transcriptional corepressor that binds to a number of
CC transcription factors. Inhibits the transcriptional activation
CC mediated by CTNMB1 and TCF family members in Wnt signaling. The
CC effects of full-length TLE family members may be modulated by
CC association with dominant-negative AES (By similarity).
CC -I- SUBUNIT: Homooligomer and heterooligomer with other family
CC members. Binds LEF1, TCF7, TCF7L1, TCF7L2 and FOXA2 (By
CC similarity).
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- TISSUE SPECIFICITY: Highly expressed in adrenal gland, small
CC intestine, kidney, lung, ovary and thyroid. Detected at lower
CC levels in pituitary, hippocampus, cortex, cerebellum and testis.
CC -I- INDUCTION: By kainic acid in the dentate gyrus.
CC -I- SIMILARITY: Contains 7 WD repeats.
CC -I- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT
CC PROTEINS.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; AF166092; AAF75590.1; -
CC InterPro; IPR005617; TLE N.
CC InterPro; IPR001680; WD40.
CC Pfam; PF03920; TLE N. 1.
CC Pfam; PF00400; WD40; 6.
CC ProDom; PD000018; WD40; 1.
CC SMART; SM00320; WD40; 7.
CC PROSITE; PS00678; WD REPEATS 1; 2.
CC PROSITE; PS50082; WD REPEATS 2; 2.
CC PROSITE; PS50294; WD REPEATS_REGION; 2.
KW Transcription regulation; Repressor; Nuclear protein; Repeat;
KW WD repeat; Phosphorylation; Wnt signaling pathway.
FT DOMAIN 1 131 GLN-RICH.
FT DOMAIN 132 191 GLY/PRO-RICH.
FT DOMAIN 199 268 CGN/PRO-RICH.
FT DOMAIN 269 456 SER/PRO-RICH.
FT DOMAIN 225 228 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT REPEAT 476 514 WD 1.
FT REPEAT 522 561 WD 2.
FT REPEAT 566 605 WD 3.
FT REPEAT 608 647 WD 4.
FT REPEAT 649 688 WD 5.

FT REPEAT 690 729 WD 6.
FT REPEAT 731 763 WD 7.
FT DOMAIN 398 406 POLY-ALA.
FT MOD_RES 240 240 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD_RES 259 259 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
FT MOD_RES 263 263 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
FT MOD_RES 267 267 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
SQ SEQUENCE 764 AA; 82643 MW; 7C5EAAEC9A1A1A2DB CRC64;
Query Match 18.9%; Score 57.5; DB 1; Length 764;
Best Local Similarity 38.3%; Pred. No. 17;
Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;
QY 2 STKDPNLDVSVSKGSGASPRITSLCTPGCKTGALMGCMKTAT 48
Db 267 SPPEGLDKARGLKADPTSPASVSSSTPSKTKDL-GHNDKST 312
RESULT 14
TLE3_MOUSE STANDARD; PRT; 771 AA.
AC 008122; Q92344;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transducin-like enhancer protein 3 (ESG) (Grg-3).
GN TLE3 OR ESG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=CD-1;
RX MEDLINE=93373944; PubMed=8365415;
RA Miyasaka H., Choudhury B.K., Hou E.W., Li S.S.-L.;
RT "Molecular cloning and expression of mouse and human cDNA encoding
RT AES and ESG proteins with strong similarity to Drosophila enhancer of
RT split groch protein."
RL Eur. J. Biochem. 216:343-352(1993).
RN (2)
RP SEQUENCE OF 41-771 FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kaenaka T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamamaki I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schiraldi L.M., Kanapin A., Matsuda H.A., Batalov S., Beisel K.W.,
RA Blake J.A., Brad D., Brusic V., Chochia C., Corbani L.B., Cousins S.,
RA Dalla B., Dragan T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaesteland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
RA Grimmer S., Guefincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan M.J., Petlea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shmida K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner U., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yangisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Tsukikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Maunishi A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yaeunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).

[3]
 RP SEQUENCE OF 323-771 FROM N.A. (ISOFORM 2).
 RC TISSUE=Breast tumor;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buelow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopfins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stacion M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP INTERACTIONS WITH LEFT1, TCF7, TCF7L1 AND TCF7L2.
 RX MEDLINE=21169341; PubMed=1126540;
 RA Brantjes H., Roose J., van De Wetering M., Clevers H.;
 RT "All Tcf Hwg box transcription factors interact with Groucho-related
 co-repressors.";
 RL Nucleic Acids Res. 29:1410-1419(2001).
 CC -1- FUNCTION: Transcriptional corepressor that binds to a number of
 CC transcription factors. Inhibits the transcriptional activation
 CC mediated by CTNNB1 and TCF family members in Wnt signaling. The
 CC effects of full-length TLE family members may be modulated by
 CC an important role during spermatogenesis.
 CC -1- SUBUNIT: Homooligomer and heterooligomer with other family
 CC members. Binds FOXA2 (by similarity). Binds LEFT1, TCF7, TCF7L1 and
 CC TCF7L2.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q08122-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q08122-2; Sequence=VSP_007025;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Expressed only in testis.
 CC -1- SIMILARITY: Contains 7 WD repeats.
 CC -1- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X73360; CAA51770.1; -;
 CC EMBL: AK031323; BAC27347.1; -;
 CC EMBL: BC006672; AA06672.1; -;
 CC PIR: S35681; S35681.
 CC MGI: 104634; TLE3.
 CC InterPro: IPR005617; TLE N.
 CC InterPro: IPR001680; WD40.
 CC Pfam: PF03920; TLE N. 1.
 CC Pfam: PF00400; WD40; 6.
 CC ProDom: PD000018; WD40; 1.
 CC PROSITE: PS00678; WD_REPEATS_1; 2.
 CC PROSITE: PS50082; WD_REPEATS_2; 2.

DR PROSITE; PS50294; WD_REPEATS_REGION; 2.
 KW Spermatogenesis; Transcription regulation; Repressor; Nuclear protein;
 KW Repeat; WD repeat; Phosphorylation; Wnt signaling pathway;
 KW Alternative splicing.
 FT DOMAIN 1 130 GLN-RICH.
 FT DOMAIN 131 197 GLY/PRO-RICH.
 FT DOMAIN 198 267 CCN DOMAIN.
 FT DOMAIN 268 450 SER/PRO-RICH.
 FT DOMAIN 224 227 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT REPEAT 483 521 WD 1.
 FT REPEAT 529 568 WD 2.
 FT REPEAT 573 612 WD 3.
 FT REPEAT 615 654 WD 4.
 FT REPEAT 656 695 WD 5.
 FT REPEAT 697 736 WD 6.
 FT REPEAT 738 770 WD 7.
 FT DOMAIN 400 408 POLY-ALA.
 FT MOD_RES 239 239 PHOSPHORYLATION (BY CK2) (POTENTIAL).
 FT MOD_RES 258 258 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
 FT MOD_RES 262 262 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
 FT MOD_RES 266 266 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
 FT VARSPLC 350 352 Missing (in isoform 2).
 FT CONFLICT 544 544 /FTId=VSP_007025.
 FT CONFLICT 558 558 R -> G (IN REF. 1).
 FT CONFLICT 558 558 W -> R (IN REF. 2).
 SQ SEQUENCE 771 AA; 83212 MW; F82638A418F757A CRC64;
 Query Match 18.9%; Score 57.5; DB 1; Length 771;
 Beet Local Similarity 38.3%; Pred. No. 17;
 Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;
 Oy 2 STKDPNPLVSVSKDGSAPRITSLSCTPGCKTGALMGCMKAT 48
 Db 266 SPENGLKAKGKDKDAPTSPASVASSSTSPSKTKDL-GHNDAST 311
 RESULT 15
 ENTK_MOUSE STANDARD; PRT; 1069 AA.
 AC P97435.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Enteropепidase (EC 3.4.21.9) (Enterokinase).
 GN PRS87 OR ENTK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 RN NCBI_Taxid=10090;
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Duodenum;
 RX MEDLINE=98147142; PubMed=9486188;
 RA Yuan X., Zheng X., Lu D., Rubin D.C., Pung C.Y.M., Sadler J.E.;
 RT "Structure of murine enterokinase (enteropепidase) and expression in
 RT small intestine during development.";
 RL Am. J. Physiol. 274:G342-G349(1998).
 CC -1- FUNCTION: RESPONSIBLE FOR INITIATING ACTIVATION OF PANCREATIC
 CC PROTEOLYTIC PROENZYMES (TRYPSIN, CHYMOTRYPSIN AND CARBOXYPEPTIDASE
 CC A). IT CATALYZES THE CONVERSION OF TRYPSINOGEN TO TRYPSIN WHICH IN
 CC TURN ACTIVATES OTHER PROENZYMES INCLUDING CHYMOTRYPSINOGEN,
 CC CARBOXYPEPTIDASES, AND PROELASTASES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of 6-Lys-Ile-7 bond in
 CC trypsinogen.
 CC -1- SUBUNIT: HETERODIMER OF A CATALYTIC (LIGHT) CHAIN AND A
 CC MULTIDOMAIN (HEAVY) CHAIN LINKED BY A DISULFIDE BOND (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
 CC -1- PM: THE CHAINS ARE DERIVED FROM A SINGLE PRECURSOR THAT IS
 CC CLEAVED BY A TRYPSIN-LIKE PROTEASE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 2 LDL-receptor class A domains.
 CC -1- SIMILARITY: Contains 2 CUB domains.

CC -1- SIMILARITY: Contains 1 SEA domain.
 CC -1- SIMILARITY: Contains 1 SRCR domain.
 CC -1- SIMILARITY: Contains 1 MAM domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U73378; AAB37317.1; -
 DR HSSP: 007954; 1CR8.
 DR MEROPS: S01.156; -
 DR MGD: MGI.1197523; Prrs67.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000859; CUB domain.
 DR InterPro: IPR002172; LDL_receptor_A.
 DR InterPro: IPR000998; MAM domain.
 DR InterPro: IPR000082; SEA domain.
 DR InterPro: IPR001254; Ser_protase_Try.
 DR InterPro: IPR001190; Sitr_receptor.
 DR Pfam: PF00431; CUB; 2.
 DR Pfam: PF00057; 1d1_recept_a; 2.
 DR Pfam: PF00629; MAM; 1.
 DR Pfam: PF01390; SEA; 1.
 DR Pfam: PF00530; SRCR; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00261; LDLRECEPTOR.
 DR PRINTS: PR00020; MAMDOMAIN.
 DR SMART: SM00042; CUB; 2.
 DR SMART: SM00192; LDLA; 2.
 DR SMART: SM00137; MAM; 1.
 DR SMART: SM00200; SEA; 1.
 DR SMART: SM00202; SR; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS01209; LDLA_1; 2.
 DR PROSITE: PS50068; LDLA_2; 2.
 DR PROSITE: PS00740; MAM_1; 1.
 DR PROSITE: PS50060; MAM_2; 1.
 DR PROSITE: PS50024; SEA; 1.
 DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
 DR PROSITE: PS50287; SRCR_2; 1.
 DR PROSITE: PS50240; TRYPsin_DOM; 1.
 DR PROSITE: PS00134; TRYPsin_HIS; 1.
 DR PROSITE: PS00135; TRYPsin_SER; 1.
 KW Signal anchor: Glycoprotein; Hydrolase; Serine protease; Zymogen;
 KW Transmembrane; Repeat.
 FT CHAIN 1 829 NON-CATALYTIC CHAIN (HEAVY CHAIN).
 FT CHAIN 830 1069 CATALYTIC CHAIN (LIGHT CHAIN).
 FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 19 47 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 48 1069 SEA.
 FT DOMAIN 52 169 LDL-RECEPTOR CLASS A 1.
 FT DOMAIN 227 268 CUB 1.
 FT DOMAIN 270 379 MAM.
 FT DOMAIN 387 549 CUB 2.
 FT DOMAIN 569 679 LDL-RECEPTOR CLASS A 2.
 FT DOMAIN 686 724 SRCR.
 FT DOMAIN 723 816 SERINE PROTEASE.
 FT DOMAIN 830 1069 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 874 874 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 925 925 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 1021 1021 BY SIMILARITY.
 FT DISULFID 229 242 BY SIMILARITY.
 FT DISULFID 236 255 BY SIMILARITY.
 FT DISULFID 249 266 BY SIMILARITY.
 FT DISULFID 688 700 BY SIMILARITY.
 FT DISULFID 695 713 BY SIMILARITY.

FT DISULFID 707 722 BY SIMILARITY.
 FT DISULFID 817 945 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 859 875 BY SIMILARITY.
 FT DISULFID 959 1027 BY SIMILARITY.
 FT DISULFID 991 1006 BY SIMILARITY.
 FT DISULFID 1017 1045 BY SIMILARITY.
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 675 675 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 751 751 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 770 770 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 791 791 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 897 897 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 936 936 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 999 999 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1069 AA; 118735 MW; E625495463743C3D CRC64;

Query Match 18.9%; Score 57.5; DB 1; Length 1069;
 Best Local Similarity 39.3%; Pred. No. 24;
 Matches 24; Conservative 6; Mismatches 24; Indels 7; Gaps 4;

QY 2 STKQFNIDL-VSVSKSGASPRITSL---CTPGCKTALMGCKNKTKAT-CHCSIHV 55
 DB 167 SLSDPTAVPVTTSDKLTSSPMFTTSASLGNLSVVAATTSAPL-CNISTATFATTSCHV 225

QY 56 S 56
 DB 226 S 226

Search completed: January 12, 2004, 14:24:51
 Job time : 20 sec

THIS PAGE BLANK (USPTO)


```

DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Lanthibiotic ericin Sa.
GN ERISA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Al3.
RA Stein T., Borchert S., Conrad B., Feesche J., Entian K.-D.,
RA Hofemeister J.;
RT "A subtilin-like gene cluster of Bacillus subtilis Al3 encodes two
RT (putative) lanthibiotics, ericin A and ericin S.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF233755; AAL15567.1; -.
DR InterPro; IPR006079; Lan.dom.
DR Pfam; PF02052; Gallidermin; 1.
DR PRINTS; PR00324; NISIN.
SQ SEQUENCE 56 AA; 6195 MW; D6F1E08B8C407D1 CRC64;

Query Match 41.1%; Score 125; DB 2; Length 56;
Best Local Similarity 60.5%; Pred. No. 6.7e-09;
Matches 23; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 5 DPNLDVSVSKDGSASPRITSLCTPGCKTGALMGC 42
DB 9 DFDDLVVKVSKDGSKITPQVLSKSLCTPGCITGPIQTC 46

RESULT 3
OQNWVS ID Q8NVW5 PRELIMINARY; PRT; 47 AA.
AC Q8NVW5;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE BsaA2 protein.
GN BSA2 OR MW1765.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004828; BAB95630.1; -.
DR InterPro; IPR006078; Gallidermin.
DR InterPro; IPR006079; Lan.dom.
DR Pfam; PF02052; Gallidermin; 1.
DR PRINTS; PR00323; GALLIDERMIN.
KW Complete proteome.
SQ SEQUENCE 47 AA; 5008 MW; 14715E32B0413532 CRC64;

Query Match 25.3%; Score 77; DB 16; Length 47;
Best Local Similarity 48.7%; Pred. No. 0.01;
Matches 19; Conservative 5; Mismatches 11; Indels 4; Gaps 2;

OY 4 KDFNLDL--VSVSKDGSASPRITSLCTPGC-KTGA 38
DB 3 KYLDLDVQKANNNSDASGDRITSHSLCTPGCAKTS 41

RESULT 4
O91V17 ID Q91V17 PRELIMINARY; PRT; 456 AA.

```

```

AC Q91V17;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 49.8 kDa protein.
GN RNHL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strusberg R.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010331; AAH10331.1; -.
DR MGI; MGI:1195456; Rn1l.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RN1lh.
DR InterPro; IPR003590; LRR_RN1lh_sub.
DR Pfam; PF00560; LRR_3.
DR SMART; SM00368; LRR_R1; 4.
DR PROSITE; PS05053; LRR_R1; 4.
SQ SEQUENCE 456 AA; 49816 MW; 007B782F05A357B8 CRC64;

Query Match 22.0%; Score 67; DB 11; Length 456;
Best Local Similarity 38.0%; Pred. No. 2.2;
Matches 19; Conservative 7; Mismatches 18; Indels 6; Gaps 3;

OY 8 LDVSVSKDGSASPRITSLCTPGCKTGAL--MGNNKATATC--HCSI 53
DB 284 LSIASNELKDEGA--RLICEBLLBRCQLBSIMIKTCSLTASCPYFCSV 331

RESULT 5
O8WUL3 ID Q8WUL3 PRELIMINARY; PRT; 567 AA.
AC Q8WUL3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to MEGF10 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strusberg R.;
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020198; AAH20198.1; -.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF; 8.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00180; EGF_Lam; 4.
DR PROSITE; PS00022; EGF_1; 10.
DR PROSITE; PS01186; EGF_2; 10.
DR EGF-like domain; Laminin EGF-like domain.
SQ SEQUENCE 567 AA; 60797 MW; CF2FB8CDB7CF627 CRC64;

Query Match 20.7%; Score 63; DB 4; Length 567;
Best Local Similarity 53.8%; Pred. No. 9.2;
Matches 14; Conservative 1; Mismatches 7; Indels 4; Gaps 2;

OY 27 TSLCTPGCKTGALMGNNKATATCHCS 52
DB 149 TSKCO--CKNGAL--CNPITGACHCA 170

RESULT 6

```

096K67
ID 096K67 PRELIMINARY: PRT: 1140 AA.
AC 096K67
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE MEGF10 protein (Hypoetical protein KIAA1780).
GN MEGF10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 9:85-95(2001).
DR EMBL, AB058676; BAB47409.1; -;
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF_14.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00180; EGF_Lam; 6.
DR PROSITE; PS00022; EGF_1; 17.
DR PROSITE; PS01186; EGF_2; 17.
KW Hypothetical protein; EGF-like domain; Laminin EGF-like domain.
SQ SEQUENCE 1140 AA; 123204 MW; 45B2FA239423895A CRC64;

Query Match 20.7%; Score 63; DB 4; Length 1140;
Best Local Similarity 53.8%; Pred. No. 19;
Matches 14; Conservative 1; Mismatches 7; Indels 4; Gaps 2;

Oy 27 TSLCTPGCKTGALMGCMKMTATCHCS 52
Db 149 TSRQO-CRNGAL-CNPITGACHCA 170
RESULT 7
O9FKZ9
ID 09FKZ9 PRELIMINARY: PRT: 752 AA.
AC 09FKZ9
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Gb|AA871479.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneo T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned PI and TAC clones."
RL DNA Res. 5:131-145(1998).
DR EMBL; AB010700; BAB08623.1; -;
DR InterPro; IPR006595; CTLH.
DR SMART; SM00668; CTLH; 2.
SQ SEQUENCE 752 AA; 84436 MW; 1FE23D5DC461AFC2 CRC64;

Query Match 20.4%; Score 62; DB 10; Length 752;
Best Local Similarity 37.7%; Pred. No. 17;
Matches 20; Conservative 6; Mismatches 21; Indels 6; Gaps 2;

Oy 2 STKDFNLIV----SVSKDSGA--SPRINTSLCTPGCKTGALMGCMKMTAT 48
Db 652 SLKEVNTDLIGATSKSKKDSNTNLSSQVTTSSSTMTSEDCSSSLMMMTQT 704

RESULT 8
ID 08G572 PRELIMINARY: PRT: 475 AA.
AC 08G572
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Deoxyguanosinetriphosphate triphosphohydrolase.
GN DGT OR B1148.
OS Bifidobacterium longum
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787;
RA Schell M.A., Karmilantzon M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA Fridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AE014738; AAN24953.1; -;
KW Hydrolase; Complete proteome.
SQ SEQUENCE 475 AA; 52701 MW; 2CF00D99B2F8C907 CRC64;

Query Match 20.2%; Score 61.5; DB 16; Length 475;
Best Local Similarity 31.1%; Pred. No. 12;
Matches 14; Conservative 8; Mismatches 22; Indels 1; Gaps 1;

Oy 11 VSVSKDSGASPRITST-SLCTPGCKTGALMGCMKMTATCHCS 54
Db 106 ILIAGTDPARTRLHTLEVAQIGRQIGALGCDPVVDCACLAH 150
RESULT 9
O8TBS6
ID 08TBS6 PRELIMINARY: PRT: 536 AA.
AC 08TBS6
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 2610027L16 gene.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Lung;
RX Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025332; AAH25332.1; -;
DR InterPro; IPR001313; Pumilio/Puf.
SQ SEQUENCE 536 AA; 58248 MW; 408722C248F2851C CRC64;

Query Match 20.2%; Score 61.5; DB 4; Length 536;
Best Local Similarity 28.3%; Pred. No. 14;
Matches 13; Conservative 9; Mismatches 23; Indels 1; Gaps 1;

Oy 11 VSVSKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCH-CS 55
Db 100 LSTNRGSEMLQELGFSPLKPLCRVWALRSLNRTVACHRCGVHV 145
RESULT 10
Q8IVF0

ID Q81VE0 PRELIMINARY; PRT; 648 AA.
 AC Q81VE0;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein KIAA2021 (Fragment).
 GN KIAA2021.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Nagase T., Kikuno R., Ohara O.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human."
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBD databases.
 DR EMBL; AB095941; BAC23117.1; -
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 648 AA; 70727 MW; D43E7A68A4EBE05 CRC64;
 Query Match 20.1%; Score 61.5; DB 4; Length 648;
 Best Local Similarity 28.3%; Pred. No. 17;
 Matches 13; Conservative 9; Mismatches 23; Indels 1; Gaps 1;
 Oy 11 VSVSKDGSASPRITSLCTPGCKTGALMGCMKATGCH-CSIHV 55
 Db 112 LSTNRIGSEMLQELGFLKPLCKRWALRSNRIVACHRCGVHV 157
 RESULT 11
 09V5J7 PRELIMINARY; PRT; 1350 AA.
 AC 09V5J7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE CG12908 protein.
 GN CG12908.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelinkker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Milos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Bailew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Bocchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis J.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dietz S.M., Dunn P.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Ibeagwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodita C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laesle P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusaker D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Retter K., Remington K., Saunders R.D.C., Scheider F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E.C., Spidling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195 (2000).
 RN (12)
 RP SEQUENCE FROM N.A.
 RA Gelinkker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fartan D.,
 RA Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibeagwam C., Jalali M., Kruse D., Li P., Matei B., Moshrefi A., J.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phoumanavong S., Pittman G.S., Puri V., Richards S., Scheider F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBD databases.
 RN (13)
 RP SEQUENCE FROM N.A.
 RA Mistra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Gelinkker S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Seale S.M.J., Smith E., Shu S., Smurciak P., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBD databases.
 RN (14)
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Gelinkker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBD databases.
 RN (15)
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBD databases.
 DR EMBL; AB003830; AAF58809.3; -
 DR FlyBase; FBgn0033509; CG12908.
 DR InterPro; IPR00152; Aa_hydroxyl.
 DR InterPro; IPR00181; EGF_Ca.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000033; Icd_receptor_rep.
 DR InterPro; IPR003886; Nidogen_ext.
 DR Pfam; PF00008; EGF_6.
 DR Pfam; PF00058; Icd_recept_b; 2.
 DR SMART; SM00181; EGF_11.
 DR SMART; SM00135; LY; 4.
 DR SMART; SM00539; NIDO; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 2.
 DR PROSITE; PS01186; EGF_2; 9.
 DR PROSITE; PS01187; EGF_Ca; 1.
 KW EGF-like domain.
 SO SEQUENCE 1350 AA; 149081 MW; 482243B79347A341 CRC64;
 Query Match 20.1%; Score 61; DB 5; Length 1350;
 Best Local Similarity 38.7%; Pred. No. 42;
 Matches 12; Conservative 3; Mismatches 12; Indels 4; Gaps 1;

Qy 30 CTGCKTGALMGCMKATATC-----CSIHVS 56
Db 940 CLDGVGQDALTCGTSKPLSCHVANNCGIHAT 970

RESULT 12

0924P4 PRELIMINARY; PRT; 456 AA.
AC 0924P4.
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Ribonuclease/angiogenesis inhibitor.
GN RNH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
[1]
RP SEQUENCE FROM N.A.
RA Meinel M.B., Comb M.J.;
RT "Mouse homolog of ribonuclease/angiogenesis inhibitor";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF071546; AAK68859.1; --
DR MGI; MGI:1195456; Rnh1.
DR InterPro; IPR001611; LRR_RNinh.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR Pfam; PF00560; LRR_3.
DR SMART; SM00368; LRR_R1; 1.
DR PROSITE; PSS0503; LRR_R1; 4.
SQ SEQUENCE 456 AA; 45626 MW; 23418247194604E5 CRC64;

Query Match 19.7%; Score 60; DB 11; Length 456;
Best Local Similarity 36.0%; Pred. No. 18;
Matches 18; Conservative 7; Mismatches 19; Indels 6; Gaps 3;

Qy 8 LDIYSVSKSDGASPRITSTSLCTPGCKTGMAL--MCGMKATATC--HCSI 53
Db 284 LSLASNEIKDGA--RLCESLLEPXCQLESIMTKCSLTASCPYCSV 331

RESULT 13

08T919 PRELIMINARY; PRT; 773 AA.
AC 08T919.
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE AT05602P.
GN CG14982.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
[1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.U., Nunoo U., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.W.,
RA Ceiniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY075166; AAL68036.1; --
DR FLYBase; FBgn0035477; CG14982.
SQ SEQUENCE 773 AA; 84788 MW; 08046811D501CDF CRC64;

Query Match 19.7%; Score 60; DB 5; Length 773;
Best Local Similarity 34.8%; Pred. No. 31;
Matches 16; Conservative 10; Mismatches 16; Indels 4; Gaps 2;

Qy 10 LVSYSKSDGASPRITSTSLC--TP-GCKTGALMGCMKATATC 51
Db 457 LVATRRDSSGSGSHANSYCGYTPAGDYSWGSGRNTCLDRC 502

RESULT 14

09VZK7 PRELIMINARY; PRT; 773 AA.
AC 09VZK7.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG14982 protein.
GN CG14982.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
[1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkeley;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballew R.M., Baer A.V., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Churry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosser C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod W.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusseken D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RA Ceiniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Branton R.C., Rogers Y.,
RA Banno C.A., An H., Baldwin D., Bazon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferrera S., Friese E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibbegam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,

RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svitskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Mitra S., Crosby M.A., Mathews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celisner S.E.,
 RA Clamp M., Drysdale R., Emmert D., Friese E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Seattle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams W.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003479; AAF47814.2; -
 SQ SEQUENCE FBgn0035477; CG14982.
 RA FlyBase; 773 AA; 84832 MW; EID9789DE2AD912C CRC64;

Query Match 19.7%; Score 60; DB 5; Length 773;
 Best Local Similarity 34.8%; Pred. No. 31;
 Matches 16; Conservative 10; Mismatches 16; Indels 4; Gaps 2;

QY 10 LVSVKSGKSGASPRITSTSLC--TP-GCKTGLMGNMKTATCHC 51
 DB 457 LVAAIRRDGSSSTQHSANSYCGVTPAGDYSGMGGRNTECDRC 502

RESULT 15
 QP860 ID PRELIMINARY; PRT; 78 AA.
 AC QP860; 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical protein XF2284.
 GN XF2284.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxId=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=985C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitejima U.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miranda E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A., Jr., Nodrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pequeiro J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silva U.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H.,
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Melandris J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa."
 RL Nature 406:151-159 (2000).
 SQ SEQUENCE 78 AA; 8776 MW; 1CDD20E677EB32PB CRC64;

Query Match 19.6%; Score 59.5; DB 16; Length 78;
 Best Local Similarity 41.2%; Pred. No. 3.3;
 Matches 14; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

QY 23 RITSTSLCTPGCK-TGALMGNMKTATCHCSIHV 55
 DB 15 RPDVSVLTGCKQAGATLTLRVATLYCSTHI 48

Search completed: January 12, 2004, 14:25:56
 Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 11:24:26 ; Search time 3036 Seconds

(without alignments)
10335.205 Million cell updates/sec

Title: US-10-082-618-8

Perfect score: 767

Sequence: 1 agtcagcaataatcctaataa.....tcgcaagaataatcgcgcgt 767

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl1.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_ph.*
24: em_pat.*
25: em_pl.*
26: em_ro.*
27: em_scs.*
28: em_un.*
29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rod.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_hcg.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	767	100.0	767	1 STRSPAN	J04057 S.lactis an
2	767	100.0	2467	1 AF465351	AF465351 Lactococc
3	765.4	99.8	1575	1 AF420259	AF420259 Lactococc
4	765.4	99.8	2689	1 LACSPAN	D00696 Lactococcus
5	750.2	97.8	7500	1 L1NISI	X68307 L.lactis ge
6	687.4	89.6	7610	1 L1NISEN	Z18947 Lactococcus
7	675	88.0	7454	6 AR228341	AR228341 Sequence
8	643	83.8	7423	1 LACNISABTC	L16226 Lactococcus
9	643	83.8	7423	6 AX002982	AX002982 Sequence
10	604	78.7	2521	1 LACNISINA	M79445 L.lactis OR
11	518	67.5	518	6 BD166172	BD166172 Leader se
12	518	67.5	518	6 I09640	I09640 Sequence 3
13	503.2	65.6	796	1 AB083093	AB083093 Lactococc
14	470	61.3	3289	1 LACNISIN	M65089 Lactococcus
15	452.6	59.0	2778	1 L1NISI	Y13384 Lactococcus
16	448	58.4	448	6 A76045	A76045 Sequence 1
17	410.4	53.5	1446	6 AX322999	AX322999 Sequence
18	400.4	52.2	402	1 LACSPANM	D10768 L.lactis mu
19	380	49.5	380	1 LACNISA	M27277 Lactococcus
20	358.4	46.7	360	1 L1NISI	X61144 L.lactis ni
21	358.4	46.7	360	6 A30280	A30280 L.lactis NI
22	358.4	46.7	360	6 AX002993	AX002993 Sequence
23	358.4	46.7	360	6 I33930	I33930 Sequence 1
24	300	39.1	300	6 A51037	A51037 Sequence 18
25	300	39.1	300	6 AR228335	AR228335 Sequence
26	295.8	38.7	300	6 AR228337	AR228337 Sequence
27	295.8	38.3	300	6 AR228336	AR228336 Sequence
28	293.6	38.3	300	6 AR228338	AR228338 Sequence
29	293.6	38.3	300	6 AR228339	AR228339 Sequence
30	290.4	37.9	300	6 AR228340	AR228340 Sequence
31	221.4	28.9	223	1 STRNISA	M24527 Streptococc
32	202.4	26.4	212	6 AX280058	AX280058 Sequence
33	189	24.6	2982	6 AX002984	AX002984 Sequence
34	174	22.7	174	6 AX002983	AX002983 Sequence
35	84.8	11.1	546	6 A76046	A76046 Sequence 2
36	81.8	10.7	144	6 AR030515	AR030515 Sequence
37	81.8	10.7	144	6 AR119687	AR119687 Sequence
38	80	10.4	878	6 AX416008	AX416008 Sequence
39	77.8	10.1	930	6 AX415244	AX415244 Sequence
40	77.8	10.1	960	6 AX413559	AX413559 Sequence
41	77.8	10.1	249050	1 AL596165	AL596165 Listeria
42	77.8	10.1	349980	6 AX417038	AX417038 Sequence
43	77.8	10.1	349980	6 AX417041	AX417041 Sequence
44	62.2	8.1	110000	2 PFMAL13_09	PFMAL13_09 Cont
45	60.6	7.9	66993	2 AC138074	AC138074 Homo sapi

ALIGNMENTS

RESULT 1
STRSPAN 767 bp DNA linear BCT 15-FEB-1996
LOCUS S.lactis antibiotic nisin (span) gene, complete cds.
DEFINITION J04057.1 GI:153816
VERSION J04057.1
KEYWORDS nisin, small protein antibiotic.
SOURCE Lactococcus lactis
ORGANISM Lactococcus lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 767)
AUTHORS Buchman,G.W., Banerjee,S. and Hansen,J.N.
TITLE Structure, expression, and evolution of a gene encoding the precursor of nisin, a small protein antibiotic

Pred. No. is the number of results predicted by chance to have a

CDS

1364..>2467
 /gene="nlsB"
 /codon_start=1
 /transl_table=11
 /product="NlsB"
 /protein_id="AAL73519.1"
 /db_xref="GI:18481690"
 /translation="MKSFFKAOPFLVNTILSPNDRKSFTEYQVIEVSKNVFLE
 QLLANPKLYDMQKYNAGLKKRVKLFESIKYKRSYLRSTPGLFSEISIGVF
 SKSROYKLMGKTKGIRLDIOWLIRLVHKMEVDPSKLSFRNNANVKGDRVQVYT
 INSELEEVNIKYNVYQIISFCEVDYQKEDICEVTLCEGDSEYSEIIGEGI
 VNHVILSNLQMDLSDFSMNFPLIVEAIDEDKXIIPLKRVQKFIQESYSEIIGEGI
 EKLKEIYQMSQILENDNYIQIDILSDSEINVDVOKOOLEHLEFLNNTKSVARTY
 LDYVKKDRIEKKGVDOEVQITELFSDSTGIGAVD"

BASE COUNT 913 a 317 c 412 g 825 t

ORIGIN

Query Match 100.0%; Score 767; DB 1; Length 2467;
 Best Local Similarity 100.0%; Pred. No. 9.8e-123;
 Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGTTGACGAATTTTAAATTTTAAATTTTAAATTTTCTAGTTCTCGAATATATA 60
 |||||
 786 AGTTGACGAATTTTAAATTTTAAATTTTAAATTTTCTAGTTCTCGAATATATA 845
 |||||
 61 GAGATAGGTTTATGAGCTTGAACATCTGAAGACCTGATTAATTAATTAATTAAT 120
 |||||
 846 GAGATAGGTTTATGAGCTTGAACATCTGAAGACCTGATTAATTAATTAATTAAT 905
 |||||
 121 AATAGAAACATTAAACAATCTTAAACAGTCTTAATTTCTTGAAGAAAGTATGTAAT 180
 |||||
 906 AATAGAAACATTAAACAATCTTAAACAGTCTTAATTTCTTGAAGAAAGTATGTAAT 965
 |||||
 181 AATATTATTGTCGATTAACGCGACATTAACGCGCTGATTAATTTCTGAAGTTTGT 240
 |||||
 966 AATATTATTGTCGATTAACGCGACATTAACGCGCTGATTAATTTCTGAAGTTTGT 1025
 |||||
 241 AGATCAATGATTTGGTTCGAGAACTAACAATTAATTAAGAGGACCTCAAAATG 300
 |||||
 1026 AGATCAATGATTTGGTTCGAGAACTAACAATTAATTAAGAGGACCTCAAAATG 1085
 |||||
 301 AGTACAAAGATTTTAACTTGATTTGGTTCGAGAAAGATTCAGGTGCATCA 360
 |||||
 1086 AGTACAAAGATTTTAACTTGATTTGGTTCGAGAAAGATTCAGGTGCATCA 1145
 |||||
 361 CCAGCAATTAACAAGTATTCGCTATGTAACCCGGTGTGTAACAGAGCTTGATGGGT 420
 |||||
 1146 CCAGCAATTAACAAGTATTCGCTATGTAACCCGGTGTGTAACAGAGCTTGATGGGT 1205
 |||||
 421 TGTATCAATGAAACAGCAACTGCTATGTAATTCAGTAAAGCAATTAACCAATCA 480
 |||||
 1206 TGTATCAATGAAACAGCAACTGCTATGTAATTCAGTAAAGCAATTAACCAATCA 1265
 |||||
 481 AGGATAGTATTTTGTAGTTGAGACATGATCTATTCCTATTTTATTAAGTATTTAGG 540
 |||||
 1266 AGGATAGTATTTTGTAGTTGAGACATGATCTATTCCTATTTTATTAAGTATTTAGG 1325
 |||||
 541 TTGCTAAATAGCTTAAATTAAGAGAGAGAAACATGATTAAGTTCATTTAAG 600
 |||||
 1326 TTGCTAAATAGCTTAAATTAAGAGAGAGAAACATGATTAAGTTCATTTAAG 1385
 |||||
 601 CTCAACCGTTTTAGTAAAGAAATACATTTTATCTCCAAACGATTAAGGAGTTTACTG 660
 |||||
 1386 CTCAACCGTTTTAGTAAAGAAATACATTTTATCTCCAAACGATTAAGGAGTTTACTG 1445
 |||||
 661 AATATACTCAAGTCACTGAGACTGTAAGTAAATTAAGTTCATTTTGGAAAGTACTAG 720
 |||||
 1446 AATATACTCAAGTCACTGAGACTGTAAGTAAATTAAGTTCATTTTGGAAAGTACTAG 1505
 |||||
 721 TAGCTAATCTTAAACTCTATGATGTTATGACAGAAATATTAATGCTGCT 767
 |||||
 1506 TAGCTAATCTTAAACTCTATGATGTTATGACAGAAATATTAATGCTGCT 1552
 |||||

RESULT 3
 AF420259
 LOCUS
 DEFINITION
 Lactococcus lactis Nisz (nls2) gene, complete cds; and NlsB (nlsB) gene, partial cds.
 AF420259
 AF420259.1 GI:18656614
 SOURCE
 ORGANISM
 Lactococcus lactis
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Lactococcus.
 1 (bases 1 to 1575)
 Lee, K.-H., An, J.-Y., Yun, H.-D., Lee, H.-J., Ji, G.-E., Chang, H.-C., Chung, D.-K., Lee, J.-H. and Kim, J.-H.
 Characterization of a nisin-producing Lactococcus lactis strain isolated from Kimchi
 2 (bases 1 to 1575)
 Lee, K.-H., An, J.-Y., Yun, H.-D., Lee, H.-J., Ji, G.-E., Chang, H.-C., Chung, D.-K., Lee, J.-H. and Kim, J.-H.
 Direct Submission
 Submitted (17-SEP-2001) Division of Applied Life Science, Gyeongsang National University, 900 Gaja-dong, Jinju, Gyeongsang 660-701, Korea
 FEATURES
 source
 1..1575
 /organism="Lactococcus lactis"
 /mol_type="genomic DNA"
 /db_xref="taxon:1358"
 745..918
 /gene="nls2"
 745..918
 /gene="nls2"
 /codon_start=1
 /transl_table=11
 /product="Nls2"
 /protein_id="AAL77866.1"
 /db_xref="GI:18656615"
 /translation="MSTKDFNLDLVSVSKDSGASPRITSISLCTPGCKTGALMGNN
 KTAACNCSINHSK"
 1026..>1575
 /gene="nlsB"
 1026..>1575
 /gene="nlsB"
 /codon_start=1
 /transl_table=11
 /product="NlsB"
 /protein_id="AAL77867.1"
 /db_xref="GI:18656616"
 /translation="MKSFFKAOPFLVNTILSPNDRKSFTEYQVIEVSKNVFLE
 QLLANPKLYDMQKYNAGLKKRVKLFESIKYKRSYLRSTPGLFSEISIGVF
 SKSROYKLMGKTKGIRLDIOWLIRLVHKMEVDPSKLSFRNNANVKGDRVQVYT
 INSELEEVNIKYNVYQIISF"

BASE COUNT 530 a 198 c 257 g 530 t

ORIGIN

Query Match 99.8%; Score 765.4; DB 1; Length 1575;
 Best Local Similarity 99.9%; Pred. No. 2e-122;
 Matches 766; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 AGTTGACGAATTTTAAATTTTAAATTTTAAATTTTCTAGTTCTCGAATATATA 60
 |||||
 448 AGTTGACGAATTTTAAATTTTAAATTTTAAATTTTCTAGTTCTCGAATATATA 507
 |||||
 61 GAGATAGGTTTATGAGCTTGAACATCTGAAGACCTGATTAATTAATTAATTAAT 120
 |||||
 508 GAGATAGGTTTATGAGCTTGAACATCTGAAGACCTGATTAATTAATTAATTAAT 567
 |||||
 121 AATAGAAACATTAAACAATCTTAAACAGTCTTAATTTCTATCTGAGAAAGTATGTAAT 180
 |||||
 568 AATAGAAACATTAAACAATCTTAAACAGTCTTAATTTCTATCTGAGAAAGTATGTAAT 627
 |||||

Qy	181	AATTTATATGTCGATPAAGCGAGCTAATTAACCGGCTCGATTTAAATCTGAAGTTTGT	240
Db	628	AATTTATATGTCGATPAAGCGAGCTAATTAACCGGCTCGATTTAAATCTGAAGTTTGT	687
Qy	241	AGATPAACAATGATTTGGTTCGAAGCAATCAAAATTAATTAAGAGGCACTCAAAATG	300
Db	688	AGATPAACAATGATTTGGTTCGAAGCAATCAAAATTAATTAAGAGGCACTCAAAATG	747
Qy	301	AGTACAAAAGATTTTAACTTGGATTTGGTATCTGTTTGAAGAAAGATTCAGTGCATCA	360
Db	748	AGTACAAAAGATTTTAACTTGGATTTGGTATCTGTTTGAAGAAAGATTCAGTGCATCA	807
Qy	361	CCAGCGATTAACAAGTATTCGTATGTACACCCGGTTGTAAACGAGAGCTGTGATGGT	420
Db	808	CCAGCGATTAACAAGTATTCGTATGTACACCCGGTTGTAAACGAGAGCTGTGATGGT	867
Qy	421	TGTAAACATGAAAAACAGCAACTTGTCTATTGTAGTATTCAAGTAAAGCAATCA	480
Db	868	TGTAAACATGAAAAACAGCAACTTGTATTTGTAGTATTCAAGTAAAGCAATCA	927
Qy	481	AGGATAGTATTTTGTAGTTCAGACATGGAATCTATCTATTTTATAAGTTATTAGG	540
Db	928	AGGATAGTATTTTGTAGTTCAGACATGGAATCTATCTATTTTATAAGTTATTAGG	987
Qy	541	TTGCTAAATAGCTTATTAATAAATPAAGAGGAAAAACATATTAATAAAGTTCATTTAAG	600
Db	988	TTGCTAAATAGCTTATTAATAAATPAAGAGGAAAAACATATTAATAAAGTTCATTTAAG	1047
Qy	601	CTCAACCGTTTATAGTAAGAATAACAAATTTATCTCAACGATPAACGAGTTTACTG	660
Db	1048	CTCAACCGTTTATAGTAAGAATAACAAATTTATCTCAACGATPAACGAGTTTACTG	1107
Qy	661	AATATACTCAAGTCACTGAGACTGTAAAGTAAATAATPAAGTTTGTGAAACAGTTACTAC	720
Db	1108	AATATACTCAAGTCACTGAGACTGTAAAGTAAATAATPAAGTTTGTGAAACAGTTACTAC	1167
Qy	721	TAGCTAAATCCTAACCTATATATGTATATGCAAGAAATATAAGCTGGT	767
Db	1168	TAGCTAAATCCTAACCTATATATGTATATGCAAGAAATATAAGCTGGT	1214

```

RESULT 4
LACSPAN
LOCUS
DEFINITION LACSPAN 2689 bp DNA linear BCT 21-DEC-2002
ACCESSION D00696 D00695
VERSION D00696.1 GI:216734
KEYWORDS transposon.
SOURCE Lactococcus lactis
ORGANISM Lactococcus lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
1 (bases 1 to 2689)
Araya,T.
Unpublished
These data kindly submitted in computer readable form by: Tomoko
Araya
Nutritional Science Laboratory
Morigata Milk Industry Co., Ltd.
1-83, 5 Higashihara
Zama, Kanagawa 228
Japan
Phone: 81-462-52-3047
Fax: 81-462-52-3055.
Location/Qualifiers
1. .2689
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/strain="NCDO497"
/db_xref="taxon:1358"
71..109
repeat unit

```

Query Match	Best Local Similarity	Matches 766; Conservative	Score 765.4; DB 1; Length 2689; Pred. No. 1.0e-122; Mismatches 1; Indels 0; Gaps 0;
DB	1849	AGTGGAGCAATTTAAATTTAAATATCTGATTTTCTAGTCCCTGAATATATA	1908
QY	61	GAGATAGATTATGAGCTTACACACTTGAATGATGACCTAGTCTTAACATACTGAC	120
DB	1909	GAGATAGATTATGAGCTTAAACACACTTGAATGACCTAGTCTTAACATACTGAC	1968
QY	121	AATGAAACATTAACAATCTAAACACGCTCTAATCTTATCTGAGAAAGTATGGTAAT	180

Db	1969	AATAGAAACATTAACAATCTAAACAGCTTAAATCTTAATCTTGAGAAAGTATGGTAAAT	2028	gene	/db_xref="taxon:1358" /clone="pGEN1,pGEN2,pGEN3"
Oy	181	AATATTATTTGTCGATAACGGAGCATATTAACCGCTCTGATTAAATCTGAAGTTGTT	240	CDS	1..329 /gene="nisX" /gene="nisX" /codon_start=3 /transl_table=11 /product="transposase" /protein_id="CAA48379.1" /db_xref="SWISS-PROT:P35878" /translation="ACGKEHPQGLIVTDDGSGQYTSRYSOTLROVGAQSSMSRKN PYDNAMESFYKTLKREILINAHFETAATQELIFKTIETTYNTRKHAHSGLDYKSPD FEKYNs"
Db	2029	AATATTATTTGTCGATAACGGAGCATATTAACCGCTCTGATTAAATCTGAAGTTGTT	2088		
Oy	241	AGATACAAATGATTCGTTGCAAGGAACCTACAAATTAATATTAAGAGGCACTCAAAATG	300		
Db	2089	AGATACAAATGATTCGTTGCAAGGAACCTACAAATTAATATTAAGAGGCACTCAAAATG	2148		
Oy	301	AGTACAAAGATTTTAACTTGATTTGGTATCTGTTTCCAGAAGATTCAGGTGCATCA	360		
Db	2149	AGTACAAAGATTTTAACTTGATTTGGTATCTGTTTCCAGAAGATTCAGGTGCATCA	2208	gene	1159..1332 /gene="nisA" /gene="nisA" /codon_start=1 /transl_table=11 /product="Nisin-prepeptide" /protein_id="CAA48380.1" /db_xref="GI:44042" /db_xref="SWISS-PROT:P13068" /translation="WSTKDPLDVSYSKDDSGASPRITSISLCTPECKTGALMGCM KATCHCSIHVSK"
Oy	361	CCACGCATTAACAGTATTTGGCTATGTACACCCGGTTGTAACAGAGCTCTGATGGGT	420	CDS	1439..4420 /gene="nisB" /gene="nisB" /codon_start=1 /transl_table=11 /note="biosynthetic gene" /codon_start=1 /transl_table=11 /protein_id="CA48381.1" /db_xref="GI:44043" /db_xref="SWISS-PROT:P20103" /translation="WIKSFRKQPEPLVANTLTLCNDKRSFTETVQVLETYSKKNKYLE QULLANPLVDYMQYKNGNLKKRKKVKKLFPSIYKYSKSLRSTPGLSESTISIGF SKSQQYKLMGKTTGIGIRLDIOWMLRLVHKMEKEDVPSKLSFRRNANFKPDRFOVYT INSSELEVNIAKYTNVYOIISEPCENDOKEDICEVTLCYGEYELSEYQYLSLI VNHVYLSNLOKDIISDPENMTPLKRVAILDEDDKYYIIPLKVKOFIOEYSEIEIGESI EUKKEITYEMKQIILENDYIOIDISBSEINPDVKQKQQLBHLAEPICANTKSVRRY LDYVDKDFTEKRGVDQVQITELFDSTFTGIGAPYNNVHPRNDPYESPSTLYSEER EYLSMVEYEAAGNHNVLINDLBESHYQMDLEKSKLQGLFLNLAKERYEKDIFILG DIVGNNNLGAAGSRFALSPELTSHRTIYDVSVERENKEITSCETVLPENIRHMAN VHNTSIRMRKVLPEFTSTSHNEVLITNVIYIGIDEXKRYARDISTOBLEKFEITYSMN KTLPSNELRLPELTSTSHNEVLITNVIYIGIDEXKRYARDISTOBLEKFEITYSMN SKMTIRELIQSKETIPKEFYINGDNKTYLSQKPLDNEILSAIKSSKKKDPTELOE YFEDENIINKBEGKRVADVVPFIRTAALNBSGAFIREKRVSVREKSLPFWMLV KLYISINQNEFLSLYDLOKILVANGNLFFLRYPDPRHRLRIKCSKDLPLAAYS ILEIKRKRKRIMSTFDISIIYDQEVERYGGFATLDESAIFCDMSKIIENLFLIKL TNNDKRVQDVSILVNYVTLKCFEPQNDKKILNPLNTLSKVKVKNVKEKILEHLKILK VNNLQDQIFYDNKRFELKHAIKNLPKMIADPELQYVYSIIDSIIHVNHNRLIGIER DEKLIYTLQRLFVSEIYMK"
Db	2509	AATTAATCTCAAGTATTTGAGACTGTAAAGTAAATTAAGTTTTTTGGAAGCTTACTAC	720		
Oy	721	TAGCTAATCTCAAGTATTTGAGACTGTAAAGTAAATTAAGTTTTTTGGAAGCTTACTAC	767		
Db	2569	TAGCTAATCTCAAGTATTTGAGACTGTAAAGTAAATTAAGTTTTTTGGAAGCTTACTAC	2615		
RESULT 5					
LOCUS	LINIS1	7500 bp	DNA	linear	BCT 19-JUL-2000
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
ABSTRACT					
COMMENT					
FEATURES					
source					


```

OULLANPKLYNMQAKYNAGLLKKRKYKLFESIIYKXKYSILRSTPFGHSEISIGVF
SSSOYKMGKTKTGIRLDTOMLIRLVHMEVDSFKLSFTRNNAKYREKSDVFOYAT
INSSELEEVNITYNYVOIISEFCENDYOKYEDICTVTLCYGEVREBELSEVLSLTI
VNHYLISNTQKDLSPSWDTPLTRKBAIDEDKTYIIPKIQKQPTQESIEIISGICI
EKUKETIYOSMSQILEMDNTIQUIDLSDSISINFDVKKQOLEHAEFLGNTTSVARTY
LDVYKDKFLEKYGVQDEVOITELFDSFTGIGAPYNNHPRNDYSEBPTLYSSEER
EKYLSWYVAVKNNHVINLADLESYHOKDLSEKSELQGLEFLNLAKREKDOIIFLG
DIVGNNNLGASGRFSGALPELTSYHRTIVDSVERENKEITSCIEIVLPENIHAN
VHSTISMRKXVLPFTSTSHNEVOLNTYIGIDEXKEFPYARDISTQEVUKFYITSMYV
KULFSNEARFLYEISLDDFKGNLPMELIYRDPDIYPRIVPDEIVISPAKMKMGDVN
NMKTIIRELIQSKEIPEKFIYVGNDKVYLSQSNPLDMEILISAIKKSKRKPILBOE
YFEDENIINKQKGVADVAVVPFIRTRALGNGRAVIRKRSYVERREKLPENMYL
KYISINRQNEFLSYLPDIQKIVANLQKGLFELRYTDPKPHIRLKCSDEKDOIFLYG
ILEILRSOKNRIMSTFEDISIYDOEVERYGGDTLESAIFCADSKIIPNLITLIDK
TNMDKVDVDSILVNYLYKRCPRQDNKILANLVPSPKRYKENVNEKIEHYTLIDK
VNLGDOIIFYDKFELKAIKOLFLKMLAQDFEJOKVSIIDSIIHVNHNLIIGIER
DKEKLIYTLQRLFVSEEYMK"
/gene=.5314
/translation="nisp"
/length=3512
/codon_start=1
/product="nisp protein"
/protein_id="CA79469.1"
/db_xref="GI:473013"
/db_xref="SPTREMBL:Q48669"
/translation="MDEVKEPTSKQFNTLTLPSTLKLIFOLEKRYAIVLILNAT
AVPLASLFIYODLINSVLGSRHLIYIIEIVOVITTVGGEYSYSGFENDRLS
YAINMRIMTTSLSLSDYEADMYNIIEKTQODSYRFOCFMNIIVLISFILLS
SLFICTWNIGVAILLILYVYLSVFLRVGQLEFLIQKRSERETIYIYLLTHD
RSPFKIKLNINISNFIHFKGLKKGINDLAIARKTYENIFLDFLILNILLTFA
MLSVRAGKLLIGNVSLIQAIKINTYSGTQIMONIIYINTSLMEQLFEBLKESV
YHKIIEDEICQODIGTAVINLSYVPSNAPALKNILSFEKELAIIVENSGSK
STLVKISGLYQPMGIIQYDKRNSLAMEEFPYOKNISVLPDPFYKVELTIEENSG
DLSOSMEDEKIKVLQNLGDLKTNNOVLDLQGNQMOBESRCLSGGOWKIALART
PFKRSIYILDEPSALDPAVEKEIFDIFVALSENNISIFISHSLNAPAKANKIYVMK
DQQVEDVSHDVLRLRCQYQYBELYYSEYEDNDE"
/gene=.5295
/translation="nisc"
/length=5295
/codon_start=1
/product="nisc protein"
/protein_id="CA79470.1"
/db_xref="GI:473014"
/db_xref="SPTREMBL:Q48670"
/translation="KRIMANKKIKRKNYKIIAOWDERTKKNKENDFEGELTSTGLP
GIILMALKONSKRIYOKIKIDNYIEVLSLSTYGLTSGLYSGAGIALSILHR
BDEKYNLNDLSLNIYIEFVIEKIEGFLFNITPSPDVPVIEGIIISYLLINDBO
YDDKILITNLSNLTKEKGLISYIKSENOMSGSBMTPLGCLNGLAGLGLAGAG
CILAIAHIKYSENESLALQKIFITYEFLEIKNOPLMKOGLVADLEKKEKIVREA
SFIRDAWCGYGISLILYGLALDNDYFVDAKRIKIESAQRKIGDSVITCGYS
GIEICSLFKRLNTKPKDSYIEEFNVNSEQILLEEYGBSGTGLEGICILVLSKF
EYSINFTYRQALLFLDFPLKGGKRR"
/gene=.6548
/translation="nisp"
/length=6548
/codon_start=1
/product="nisp protein"
/protein_id="CA79465.1"
/db_xref="GI:473015"
/db_xref="SPTREMBL:Q48671"
/translation="MRKYLILVALIGITGLSGYQTSOKKVRPDEGSYTNFIYDNKS
YFTVKEIQRQENVNCKKVFNLITVDMKSEKLLSSNNGSTVTLVANNIYKASDLSLC

```

```

MGINDRYKKILPESDQGVKALRLQNPVDSIDSDNFVINDKNSRKIDYMGNTYSIS
DTTYSDEELAGEVDIAEVRVDSVSGSIRSENGRIIDKQSGNSKQSRTEWDVEIHI
SIRGSJLDELAFAVELINDPFLKATKYN"
/gene=.7287
/translation="nisp"
/length=7287
/codon_start=1
/product="nisp protein"
/protein_id="CA79466.1"
/db_xref="GI:581294"
/db_xref="SPTREMBL:Q48674"
/translation="MKIKGLFVYCSGLSATYHGETTNSQOILSNINTELINANS
NALISSTGSGTSDISINSGSIVASSTRTLELDVGAANKTLLOTSVAVOKEMKVSLOET
QVSSEF"
BASE COUNT      2802 a      945 c      1400 g      2463 t
ORIGIN
Query Match      89.6%; Score 687.4; DB 1; Length 7610;
Best Local Similarity 99.0%; Pred. No. 4e-109;
Matches 702; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
59 TAGAGATAGTTTATTTAGTCTTAGACATCTGAAAGACCTAGCTTATTAATCTACTG 118
1 TAGATATAGTTTATTTAGTCTTAGACATCTGAAAGACCTAGCTTATTAATCTACTG 60
119 ACAATAGAAACATTAACAATCTAATAACAGTCTTAATCTAATCTTGAAGAAATTTGTA 178
61 ACAATAGAAACATTAACAATCTAATAACAGTCTTAATCTAATCTTGAAGAAATTTGGA 120
179 ATAATATTATTGTCATTAACGACATTAATAACGCTCTGATTAAATTTGAAAGTTTG 238
121 ATAAATATTATTGTCATTAACGACATTAATAACGCTCTGATTAAATTTGAAAGTTTG 180
239 TTAGATCAATAGATTGTTGGAAGAACTCAAAATTAATTAAGAGGACCTCAAA 298
181 TTAGATCAATAGATTGTTGGAAGAACTCAAAATTAATTAAGAGGACCTCAAA 239
299 TGAGTACAAAGATTTTAACTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 358
240 TGAATACAAAGATTTTAACTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 299
359 CACCAAGCATTAACAAGTATTTGCTATGTACACCCGTTGTPAAAACAGAGCTTGATGG 418
300 CACCAAGCATTAACAAGTATTTGCTATGTACACCCGTTGTPAAAACAGAGCTTGATGG 359
419 GTTGAACATGAAAAACGCACTTGCTATGTACAGTTACAGTAAGCAATTAACCAATC 478
360 GTTGAACATGAAAAACGCACTTGCTATGTACAGTTACAGTAAGCAATTAACCAATC 419
479 AAAGATAGTATTTTGTAGTTAGACATGATCTATCTATTTTAAAGTTATTTAG 538
420 AAAGATAGTATTTTGTAGTTAGACATGATCTATCTATTTTAAAGTTATTTAG 479
539 GGTTCCTAATAGCTTATTAATAAATTAAGAGGAAAAACATGATTAATAAGTCAATTA 598
480 GGTTCCTAATAGCTTATTAATAAATTAAGAGGAAAAACATGATTAATAAGTCAATTA 539
599 AGCTCAACCGTTTATTAAGTAAGAAATTAATTTATCTCCAAACGATTAACGAGATTAC 658
540 AGCTCAACCGTTTATTAAGTAAGAAATTAATTTATCTCCAAACGATTAACGAGATTAC 599
659 TGAATATCTCAAGTCATTTAGACCTGTAGTAAATTAATTTTGAAGACAGTTACT 718
600 TGAATATCTCAAGTCATTTAGACCTGTAGTAAATTAATTTTGAAGACAGTTACT 659
719 ACTAGCTAATCTAATCTAATGATTTATGATGTTATGCAAGAAATTAATGCTGCT 767
660 ACTAGCTAATCTAATCTAATGATTTATGATGTTATGCAAGAAATTAATGCTGCT 708

```


RESULT 7	AR228341	7454 bp	DNA	1linear	PAT 20-DEC-2002
LOCUS	AR228341				
DEFINITION	Sequence 29 from patent US 6448034.				
ACCESSION	AR228341				
VERSION	AR228341.1	GI:27267126			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 7454)				
TITLE	Gasson,M.J. and Dodd,H.M.				
JOURNAL	Production of variant nisin				
FEATURES	Patent: US 6448034-A 29 10-SEP-2002;				
source	location/Qualifiers				
	1..7454				
	/organism="unknown"				
BASE COUNT	2741 a	920 c	1376 g	2417 t	
ORIGIN					

Query Match	88.0%;	Score 675;	DB 6;	Length 7454;
Best Local Similarity	100.0%;	Pred. No. 5.4e-107;		
Matches 675; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

RESULT 8	
LACNISABCT	LACNISABCT
LOCUS	7423 bp DNA linear BCT 14-SEP-1993
DEFINITION	<i>Lactococcus lactis</i> nisin A (nisA) and nisB, nisc, nist, and nisl genes, complete cds.
ACCESSION	I16226
VERSION	I16226.1 GI:400365
KEYWORDS	nisa gene; nisb gene; nisc gene; nisl gene; nist gene; nislA A; translocator protein.
SOURCE	<i>Lactococcus lactis</i>
ORGANISM	<i>Lactococcus lactis</i> Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; <i>Lactococcus</i> .
REFERENCE	1 (bases 1 to 7423)
AUTHORS	Klippers,O.P., Beertshuyzen,M.M., Sierens,R.J. and De Vos,W.M.
TITLE	Characterization of the nisin gene cluster nlsABCPR of <i>Lactococcus lactis</i> . Requirement of expression of the nisa and nisl genes for development of immunity
JOURNAL	Eur. J. Biochem. 216 (1), 281-291 (1993)
MEDLINE	93373937
PUBMED	7689965
COMMENT	Original source text: <i>Lactococcus lactis</i> (strain NIZO R5, Sub-species <i>lactis</i>) DNA. Location/Qualifiers
FEATURES	

FEATURES
SOURCE

gene

CDS

gene
CDS

gene

```
reactions"
/codon_start=1
/codon_table=11
/label=ORF
/protein_id="AA25190.1"
/db_xref="GI:400367"
/translation="MKSSTFKQAPFLVRNTILSPNDKRSSTETVQTIVETVSKNKVLE
OLLAPKLYDVQKNAGLKKKRVKLFESLYKKYKRYLSTPGLSPSESIQVPS
KSSQYKMGKTKTGIRLDTOMLIRLVKMEVPSKLSSTRRNANKPDRFOVYIT
INSSSEAEVNIKYTNVYOIISERCENDYOKYEDICETVYLICYDEVERLSBOYLGLI
VNHLYLSNIQKDLSPSMTPLTKYBAIDEDKKYIILPKVQKFOEYSEIEGSEI
EKKEIYQESQILENDNYIQDLISDSSEFPVQKQDLEHAFSGNKTYSRRYIT
LDYKQKFLERYGVQDEOYITELPDSFTFGAGYVNRPNQDEYESPRLYVSEER
EKLSTAYVAVAKNHNINLDDLSHOKMDLSEKSELQGLLEPLNAKEYEKIFILGAG
DIYGNNNLGGASGRFSALEPILSTSVHRTIVSVERENENKETSCEIPLPFRHNN
VWHTSIMRRKVLDPFTSTSHNEVLTNTNITIGIDSEKRYARDISTQGVLFYITSMYR
KITLFSNELPLVEISLDDKFGNLPWELIYRDPYIPLRVAFDELVISPAKKMIGRDVNV
SKRTIELLOSKEIIPREYIYVGNQNKVYSQENPLDMEILESAIIKSSKKDFLELOE
YSEDDENIKNGEGRVADVVPFIRFGALGNBERAFIREKRVSEERKLPENMELVIA
KIYLSINRONEFLVSLPIDIOKIVANLGNLPELRYTDPDHTRLRKSGDLFLVAGS
IIEILRSKRKRIMSTFDISITDQEVERGGPTTELSAIFADSKILNLTLLIYD
TNNMDVADVDSIILVNTLYLAKCFQONDKKILNLNLVSPKRVENNEKIEHYLIKLK
VNNLGOIPIFYDKAFELKHAIKLFLKMTAQDELOKVSIIISIIYHNHNNRILGIER
DEKLIYKTRNFVSEEWK"
3447. 5249
/gene="t1st"
```


CDS
3447..5249
/gene="nist"
/note="putative"
/codon_start=1
/label_table=11
/label=ORF
/product="translocator protein"
/protein_id="AA025191.1"
/db_xref="GI:400368"
/translation="MDEVKFTSKQFNNLPLPSTKILPDEKRYAIVLVNAIT
AVPLASFTYODLINSVLSGRHLNIIITFYVAVITTVLGLESYVSGKEDRLS
YSINMRKMTTSLSDYEDQADMYNIJEKVQDSTYKPFQFNAIIVLSFPIILS
SLFPIGTWNIQVAILLIVPLSLVFLRQGLLEFLIQORASREIWIYVLLTHD
FSEFKLNNISNYFHKFKLKGFINODLARKKTYENIFLPIELNIIITPA
MISVRAGKLIGNLSLQISKINTSCMTQNIYIITNLSLMEOLFEEIKRESV
VHKIEDPEICNQHICTVYINLSYVPSNAPALKNTLSSEKELTRIVKNSGK
STLVKISGLIQTPTMGIQYDQRSLSMEEFYQKNTSVLFQDFVKEITLIENTGLS
DLSQWDEGKIIVKLDNLGLDELKTNNOYVLDLQNMVQEBQHGQGMOKILART
PFKKAISYILDEPSALDPVAKEIFDYFVALSENNISIFISHSINAAKANKIYVMK
DCQVEDVGSADVLRRCQYQYQELYYSEQEDNDE"
5242..6486
/gene="nisc"
5242..6486
/gene="nisc"
/note="could encode enzyme catalyzing modification
reactions"
/codon_start=1
/label_table=11
/label=ORF
/protein_id="AA025192.1"
/db_xref="GI:400369"
/translation="NNKKIKRNVKILIAQWDETRKKNKNEPFGELTLSTGLP
MIAELKNDNSKIYOKKIDNYEIVYSKSTYGLTGSYGAAGIATLSILHEDDE
KTKNLUDSLNRYIEYFKEKIEFNENITTPDYVIBGLSILSLINDQYDIL
KLILNPSNLTKENNGLSLYIKSENQSOSESNYPGCLNMGILAGLAVGCI
YAHIKGYSNEASLSLOKIIIFYEKELERKQFLMKQGLADELKEKIVREASFI
DAMCYGPGISLILYGLALNDYFVDAKEILSQAQKRGISYVNIHQYSGGLE
NCSLPRILNTKPDSPYMEEPVNSEQLIEEYGDSSGCGFLEGISGCLIVLSKFEYSI
NCTYRQALLLPDDPLKGGKR"
6483..7220
/gene="nisi"
6483..7220
/gene="nisi"
/note="encodes a protein involved in immunity against
nisin"
/codon_start=1
/label_table=11
/label=ORF
/protein_id="AA025193.1"
/db_xref="GI:400370"
/translation="MRYVILILVALIGTGLSGCYQTSHKYRFDGSGTYNFYVNSK
YVTVKEIPEQENVNNSKVFYKLLIVDMKSEKLSSSNKNSVTLVANNIYEASDLSL
MGINDRYKILPESDKAVKALQLQNFVDTSDISDNFVINDSRKIDYMGNIISIS
DITVDEDELEGVODVLAEVRFVDSVSGSIPRSEWGRIDKQDSNSKQSTRWYDEIH
SIRKSLTAPAFVEIINDPKLATKVGK"
BASE COUNT 2729 a 913 c 1374 g 2407 t
ORIGIN

Query Match 83.8%; Score 643; DB 1; Length 7423;
Best Local Similarity 100.0%; Pred. No. 1.8e-101;
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GAAACATTAAACAATTAAGAGTCTTAATCTTGAAGAAAGTATGTAATAATA 184
DB 1 GAAACATTAAACAATTAAGAGTCTTAATCTTGAAGAAAGTATGTAATAATA 60
QY 185 TATTGTGATTAACGAGCATTAATAAGGCTGATTAATTAATTTGGAATTTGTTAGT 244
DB 61 TTTATGTGATTAACGAGCATTAATAAGGCTGATTAATTAATTTGGAATTTGTTAGT 120
QY 245 ACAATGATTTGCTTCAGAGAACTACAAATTAATTAAGAGGACACTCAAAATGAGTA 304
DB 121 ACAATGATTTGCTTCAGAGAACTACAAATTAATTAAGAGGACACTCAAAATGAGTA 180

QY 305 CAAAGATTTTAACTGGATTTGGTATCTGTTTGAAGAAAGATTGAGTGATCACCAC 364
DB 181 CAAAGATTTTAACTGGATTTGGTATCTGTTTGAAGAAAGATTGAGTGATCACCAC 240
QY 365 GCATTACAGATTTTGGTATGATACACCCGGTGTGTAACAGAGCTGTGATGGGTGTA 424
DB 241 GCATTACAGATTTTGGTATGATACACCCGGTGTGTAACAGAGCTGTGATGGGTGTA 300
QY 425 ACATGAAACAGCACTTGTGATGTAGTATTCAGTAAACCAATTAACCAATCAAGA 484
DB 301 ACATGAAACAGCACTTGTGATGTAGTATTCAGTAAACCAATTAACCAATCAAGA 360
QY 485 TAGATTTTGGTATGATGACATGATCTATCCATTTTATAGTATTTAGGGTTGC 544
DB 361 TAGATTTTGGTATGATGACATGATCTATCCATTTTATAGTATTTAGGGTTGC 420
QY 545 TAAATAGCTTATTAATAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
DB 421 TAAATAGCTTATTAATAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 605 ACCGTTTGTAGTAAAGATTAATTTATCTCCAAACGATTAACGAGTTTACTGATA 664
DB 481 ACCGTTTGTAGTAAAGATTAATTTATCTCCAAACGATTAACGAGTTTACTGATA 540
QY 665 TACTCAAGCTAGTGAAGTATTAAGTAAATTAAGTATTTTGGAGAGTACTACTAGC 724
DB 541 TACTCAAGCTAGTGAAGTATTAAGTAAATTAAGTATTTTGGAGAGTACTACTAGC 600
QY 725 TAACTCAAGCTATGATGATTTATGACAGAAATTAATTAAGTCTGCT 767
DB 601 TAACTCAAGCTATGATGATTTATGACAGAAATTAATTAAGTCTGCT 643

RESULT 9
LOCUS AX002982 7423 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 1 from Patent WO9941978.
ACCESSION AX002982
VERSION AX002982.1 GI:9926884
KEYWORDS
SOURCE Lactobacillus delbrueckii subsp. lactis
ORGANISM Lactobacillus delbrueckii subsp. lactis
Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
Lactobacillus.
REFERENCE 1
AUTHORS Flanagan,A.J. and Rolph,T.P.
TITLE Performance enhancement
JOURNAL Patent: WO 9941978-A 1 26-AUG-1999;
FLANAGAN ALISON JOHNSTONE (GB); ROLPH TIMOTHY PETER (GB)
FEATURES
source
1. 7423.
/organism="Lactobacillus delbrueckii subsp. lactis"
/mol_type="genomic DNA"
/sub_species="lactis"
/db_xref="taxon:29397"
/note="NIZO RS"
RBS 159..165
RBS 440..441
RBS 444..447
RBS 3430..3433
RBS 3435..3438
RBS 5228
RBS 5230
RBS 5232..5235
RBS 6465..6473
BASE COUNT 2729 a 913 c 1374 g 2407 t
ORIGIN

Query Match 83.8%; Score 643; DB 6; Length 7423;
Best Local Similarity 100.0%; Pred. No. 1.8e-101;
Matches 643; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GAAACATTAACAATCTAAACAGCTCTTATCTATCTTGAAGAGTATGGTAATATA 184
DB 1 GAAACATTAACAATCTAAACAGCTCTTATCTATCTTGAAGAGTATGGTAATATA 60
QY 185 TTATTTGCGATTAACGGGAGCATATATAACGGCTCTGATTAATTTCTGAAGTTTGTAGAT 244
DB 61 TTATTTGCGATTAACGGGAGCATATATAACGGCTCTGATTAATTTCTGAAGTTTGTAGAT 120
QY 245 ACAATGATTTGCTGGAAGAACTACAAATATAATATPAGAGGACGACCAAAATAGATA 304
DB 121 ACAATGATTTGCTGGAAGAACTACAAATATAATATPAGAGGACGACCAAAATAGATA 180
QY 305 CAAGAATTTTAACCTGATTTGCTGATCTGTTTCGAAGAAAGATTCAGGTGATCACCAC 364
DB 181 CAAGAATTTTAACCTGATTTGCTGATCTGTTTCGAAGAAAGATTCAGGTGATCACCAC 240
QY 365 GCATTACAAAGTATTTGCTGATGATACACCGGTTGTAACAGAGCTCTGATGGGTTGTA 424
DB 241 GCATTACAAAGTATTTGCTGATGATACACCGGTTGTAACAGAGCTCTGATGGGTTGTA 300
QY 425 ACATGAAAACAGCACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 484
DB 301 ACATGAAAACAGCACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 485 TAGTATTTTGTAGTACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
DB 361 TAGTATTTTGTAGTACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 545 TAAATGCTTATTAATAAAG 604
DB 421 TAAATGCTTATTAATAAAG 480
QY 605 ACGGTTTGTAGTACAGCACTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
DB 481 ACGGTTTGTAGTACAGCACTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 665 TACTCAAGTATGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 724
DB 541 TACTCAAGTATGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 725 TAACTCTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 767
DB 601 TAACTCTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 643
RESULT 10
LOCUS LACINISINA 2521 bp DNA linear BCT 26-APR-1993
DEFINITION L.lactis ORF1 and ORF2 (nisin) gene, complete cds and ORF3 (nisin)
ACCESSION M79445.1 GI:149451
VERSION M79445.1 GI:149451
KEYWORDS nisin.
SOURCE Lactococcus lactis
ORGANISM Lactococcus lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 2521)
Dodd,H.M., Horn,N. and Gasson,M.J.
Analysis of the genetic determinant for production of the peptide
antibiotic nisin
J. Gen. Microbiol. 136 (Pt 3), 555-566 (1990)
JOURNAL MEDLINE
PUBMED 90369041
COMMENT Original source text: Lactococcus lactis DNA.
FEATURES
source
1. .2521
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"
625. .2521
/gene="nisin"
625. .1386
CDS
gene
CDS

/gene="nisin"
/note="ORF1"
/codon_start=1
/transl_table=11
/protein_id="AA025197.1"
/db_xref="GI:149452"
/translation="MHRPDSQOVERELISERIKAVFHEHGRYGVAVRITKVLNHTGI
MTNRVGLKMLHMLGVLKAGSRVYKHYNRGASISRENLINQIFKATAPKRVLMGDM
TVIPKRGTLVAVNIDVFSRKIVGSMSSSMOOLVVDCLQACGKHPPGLVHT
DQSGYSSRSRYSTLRQVGAOSMSRKPNPYDNMMESFYRTILRELINDAHFETRAE
ATQETFKTETVYNTKRHSGLDYKSPDFEKINS"
2215. .2388
/gene="nisin"
/note="ORF2"
/codon_start=1
/transl_table=11
/protein_id="AA025198.1"
/db_xref="GI:149453"
/translation="MSTWDFNLDVSVSKDGSASPRITSISLCTPGCKTGLMGCNM
KTATCHCSIHVSX"
2496. >2521
/gene="nisin"
/note="ORF3"
/codon_start=1
/transl_table=11
/protein_id="AA025199.1"
/db_xref="GI:551876"
/translation="MKSFRK"
BASE COUNT 891 a 391 c 450 g 789 t
ORIGIN
Query Match 78.7%; Score 604; DB 1; Length 2521;
Best Local Similarity 100.0%; Pred. No. 1.1e-94; Indels 0; Gaps 0;
Matches 604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTTGACGAATATTTAATATTTAATATATCTTGATTTTCTAGTCTGAAATATA 60
DB 1918 AGTTGACGAATATTTAATATTTAATATATCTTGATTTTCTAGTCTGAAATATA 1977
QY 61 GAGATAGGTTTATTAAGCTTATGACATCTGAAAGACCTGATTAATATCTGAC 120
DB 1978 GAGATAGGTTTATTAAGCTTATGACATCTGAAAGACCTGATTAATATCTGAC 2037
QY 121 AATAGAAACATTAACAAATCTAAACAGCTTAATCTATCTGAGAAAGTATGTAAT 180
DB 2038 AATAGAAACATTAACAAATCTAAACAGCTTAATCTATCTGAGAAAGTATGTAAT 2097
QY 181 AATATTTATGTCATTAACGCGAGCATTAATAACGCGCTGATTAATCTGAAGTTGTT 240
DB 2098 AATATTTATGTCATTAACGCGAGCATTAATAACGCGCTGATTAATCTGAAGTTGTT 2157
QY 241 AGATCAATGATTTGCTTGAAGAACTACAAATATAATTTATAGAGAGCACTCAAATG 300
DB 2158 AGATCAATGATTTGCTTGAAGAACTACAAATATAATTTATAGAGAGCACTCAAATG 2217
QY 301 AGTACAAAGATTTTAACCTGATTTGATGATCTGTTTGAAGAAAGATTCAGGTGATCA 360
DB 2218 AGTACAAAGATTTTAACCTGATTTGATGATCTGTTTGAAGAAAGATTCAGGTGATCA 2277
QY 361 CCACGATTAACAAGTATTTGCTATGATACACCGGTTGTAACAGAGCTCTGATGGGT 420
DB 2278 CCACGATTAACAAGTATTTGCTATGATACACCGGTTGTAACAGAGCTCTGATGGGT 2337
QY 421 TGTAAATGAAAACGCACTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 2338 TGTAAATGAAAACGCACTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397
QY 481 AGGATAGATTTTGTATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 2398 AGGATAGATTTTGTATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2457
QY 541 TTGCTAAATAGCTTATTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

Db		2458	TTCGTAATACCTTATAAATAAGAGCGAAAAAA	CATGATATAAAAGTTCAATTAAAG	251
Oy		601	CTCA	604	
Db		2518	CTCA	2521	
<hr/>					
RESULT 11					
LOCUS	BD166172				
DEFINITION	BD166172	518 bp	DNA	linear	PRT 17-JAN-2003
	Leader sequence inducing a post-translational modification of polypeptides in bacteria, and gene therefor.				
ACCESSION	BD166172				
VERSION	BD166172.1	GI:27871984			
KEYWORDS	JP 2002191383-A/2.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 518)				
AUTHORS	Hansen,N.J.				
TITLE	Leader sequence inducing a post-translational modification of polypeptides in bacteria, and gene therefor				
JOURNAL	Patent; JP 2002191383-A 2 09-JUL-2002;				
	UNIVERSITY OF MARYLAND				
COMMENT	OS Bacillus subtilis				
	PN JP 2002191383-A/2				
	PD 09-JUL-2002				
	PR 08-NOV-2001 JP 2001343857				
	PF 05-JUL-1988 US 214959				
	PI NORMAN J HANSEN				
	PC C12N15/09,C07K14/195,C12N15/00				
	CC Leader sequence inducing a post-translational modification of CC polypeptides				
	CC in bacteria, and gene therefor				
FH	Key location/Qualifiers				
FT	CDS (298)..(471).				
FEATURES	Location/Qualifiers				
- source	1..518 /organism="unidentified"				
	/mol_type="Genomic DNA"				
	/db_xref="taxon:32644"				
BASE COUNT	187 a	187 c	89 g	168 t	
ORIGIN					
Query Match	67.5%; Score 518; DB 6; Length 518;				
Best Local Similarity	100.0%; Pred. No. 9.9e-80;				
Matches	518; Conservative	0; Mismatches	0; Indels	0; Gaps	0,
Oy	1	AGTTGACGAATTTTAATTAATTTTAATTAATTCGATTTTCTAGTCCGTAATATATA	60		
Db	1	AGTTGACGAATTTTAATTAATTTTATTAATTCGATTTTCTAGTCCGTAATATATA	60		
Oy	61	GAGATAGGTTTATGAGTCTTAGACATCATCTGAAGACCCTGATTAACATATAC	120		
Db	61	GAGATAGGTTTATGAGTCTTAGACATCATCTGAAGACCCTGATTAACATATAC	120		
Oy	121	AATGAAACATTTAACAAATCTAAAACAGCTTAATTTCTATCTTGGAAGAATTTGGTAAT	180		
Db	121	AATGAAACATTTAACAAATCTAAAACAGCTTAATTTCTATCTTGGAAGAATTTGGTAAT	180		
Oy	181	AATATTATTGTGCATPACGCCGACATTAATAACGGCTCTGATTAATTTCTGAAGTTGTT	240		
Db	181	AATATTATTGTGCATPACGCCGACATTAATAACGGCTCTGATTAATTTCTGAAGTTGTT	240		
Oy	241	AGATCAATGATTTGGTTCGAAGGAACATCAAAATTAATTAAGAGGCACTCAAATG	300		
Db	241	AGATCAATGATTTGGTTCGAAGGAACATCAAAATTAATTAAGAGGCACTCAAATG	300		
Oy	301	AGTACAAAGATTTTTTAACCTTGATTTGGTATCTGTTTTCGAAGAAAGATTCAGTGTCATCA	360		
Db	301	AGTACAAAGATTTTTTAACCTTGATTTGGTATCTGTTTTCGAAGAAAGATTCAGTGTCATCA	360		
Oy	361	CCAGCGATTAACAATTTCCGTATATACACCGCGTGTGAACAAGAGACTCTGATGGGT	420		

Db	361	CCAGGCAATTACAACTATTTCGCTATGTACACCCGGTTGTAAAAACAGAGCGCTCTGATGGGT	420
Qy	421	TGTAACTGAAAACAGCACTTGTGCTTGTGTATTTCACTGAGCAATAACCAATCAA	480
Db	421	TGTAACTGAAAACAGCACTTGTGCTTGTGTATTTCACTGAGCAATAACCAATCAA	480
Qy	481	AGGATGATATTGTTGTTAGTTCAAGCATGGAATACATCC	518
Db	481	AGGATGATATTGTTGTTAGTTCAAGCATGGAATACATCC	518
RESULT 12			
LOCUS	109640	109640	518 bp DNA linear PAT 02-DEC-1994
DEFINITION	Sequence 3 from Patent WO 9000558.		
ACCESSION	109640		
VERSION	109640.1	GI:587654	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 518)		
TITLE	Hansen, N. J.		
JOURNAL	LEADER SEQUENCE INDUCING A POST-TRANSLATIONAL MODIFICATION OF		
FEATURES	POLYPEPTIDES IN BACTERIA, AND GENE THEREFOR		
source	Patent: WO 9000558-A 3 25-JAN-1990;		
	Location/Qualifiers		
	1..518		
	/organism="unknown"		
BASE COUNT	187 a	74 c	89 g 168 t
ORIGIN			
Query Match	67.5%; Score 518; DB 6; Length 518;		
Best Local Similarity	100.0%; Pred. No. 9.9e-80;		
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	AGTTGACGAATATTTAATTAATTTTATTAATTAATCTTGATTTTCTAGTTCTGGAATATA	60
Db	1	AGTTGACGAATATTTAATTAATTTTATTAATTAATCTTGATTTTCTGGAATATA	60
Qy	61	GAGATAGTTTATGAGCTCTTACGACTTGAATGATGCTAGTCTTAACTATACATGAC	120
Db	61	GAGATAGTTTATGAGCTCTTACGACTTGAATGATGCTAGTCTTAACTATACATGAC	120
Qy	121	AATGAACAATTAACAATCTAAACAGCTTAATCTATCTTGAGAAAGTATTGTAAT	180
Db	121	AATGAACAATTAACAATCTAAACAGCTTAATCTATCTTGAGAAAGTATTGTAAT	180
Qy	181	AATATTATGTCGATAACGCGAGCATATAAACGGCTCTGATTAAATTCGAAGTTGTT	240
Db	181	AATATTATGTCGATAACGCGAGCATATAAACGGCTCTGATTAAATTCGAAGTTGTT	240
Qy	241	AGATACAAATGATTCGTTCCGAAGAACTACAAAATAATTTAAGAGGCACTAAATG	300
Db	241	AGATACAAATGATTCGTTCCGAAGAACTACAAAATAATTTAAGAGGCACTAAATG	300
Qy	301	AGTACAAAGATTTTAACTGGATTTGATTCGTCTGTTTGAAGAAAGATTCAGTGATCA	360
Db	301	AGTACAAAGATTTTAACTGGATTTGATTCGTCTGTTTGAAGAAAGATTCAGTGATCA	360
Qy	361	CCAGCAATTAACAATATTTGCTATGTACACCCGGTGTGTAAACAGAGAGCTCTGATGGGT	420
Db	361	CCAGCAATTAACAATATTTGCTATGTACACCCGGTGTGTAAACAGAGAGCTCTGATGGGT	420
Qy	421	TGTAACTGAAAACAGCACTTGTGCTTGTGTATTTCACTGAGCAATAACCAATCAA	480
Db	421	TGTAACTGAAAACAGCACTTGTGCTTGTGTATTTCACTGAGCAATAACCAATCAA	480
Qy	481	AGGATGATATTGTTGTTAGTTCAAGCATGGAATACATCC	518
Db	481	AGGATGATATTGTTGTTAGTTCAAGCATGGAATACATCC	518

RESULT 13
AB083093
LOCUS
DEFINITION
AB083093
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

AB083093 796 bp DNA linear BCT 14-MAY-2003
Lactococcus lactis subsp. lactis nisZ, nisB genes for nisin Z,
nisin B, complete and partial cds.
AB083093.1 GI:23496476
Lactococcus lactis subsp. lactis
Lactococcus lactis subsp. lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
1. Park, S.H., Itoh, K., Kikuchi, E., Niwa, H. and Fujisawa, T.
Identification and Characteristics of Nisin Z-Producing Lactococcus
lactis subsp. lactis Isolated from Kimchi
Curr. Microbiol. 46 (5), 385-388 (2003)
22617875
12732968
2. (bases 1 to 796)
Park, S.H.
Direct Submission
Submitted (02-APR-2002) Sang-Hee Park, Tokyo University, Laboratory
of Veterinary Public Health, Bunkyo-ku Yagoy 1-1-1, Tokyo 113-8657,
Japan (E-mail: park3hee@hotmail.com, Tel:81-3-5841-5476,
Fax:81-3-5841-8188)

FEATURES

source

1. 796
/organism="Lactococcus lactis subsp. lactis"
/mol_type="genomic DNA"
/isolate="K231"
/isolation_source="Kimchi"
/sub_species="lactis"
/db_xref="taxon:1360"
39. 212
/gene="nisZ"
39. 212
/gene="nisZ"
/codon_start=1
/transl_table=11
/product="nisin Z"
/protein_id="BAC20309.1"
/db_xref="GI:23496477"
/translation="MSTKDFNLDIVSVSKDSGSPRITSISLCTPGCKTALMGCM
KATNCSTHVS K"
320. 796
/gene="nisB"
320. 796
/gene="nisB"
/codon_start=1
/transl_table=11
/product="nisin B"
/protein_id="BAC20310.1"
/db_xref="GI:23496478"
/translation="MKSSEKAPFLVNTILSPNDKSEFTEYTOVIEIVSKYEL
OLLIANPKLVNMOKYVAGLKKRVKLFESIIYVRSYLRSPREGFSTSTGCV
SKSSQYKLMGKTTTGIRLDTQWLRVHMEVDFSKLSFTNNANYKGDVFOYV"

BASE COUNT 303 a 99 c 141 g 253 t
ORIGIN

Query Match 65.6%; Score 503.2; DB 1; Length 796;
Best Local Similarity 99.4%; Pred. No. 3.2e-77;
Matches 505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 260 GAGGAACACTACAAATTAATTAAGAGCACTCAAAATGAGTACAAAAGATTTTAACT 319
DB 1 GAAGAACTACAAATTAATTAAGAGCACTCAAAATGAGTACAAAAGATTTTAACT 60
QY 320 TGGATTTGGTATCTGTTTGAAGAAAGATTCAGGTGCATCACCAGCATTTACAGTATT 379
DB 61 TGGATTTGGTATCTGTTTGAAGAAAGATTCAGGTGCATCACCAGCATTTACAGTATT 120

QY 380 CGCTATGTACACCCGGTTGTAACAGAGCTCTGATGGGTTGTAACTAAGAAACAGCA 439
DB 121 CGCTATGTACACCCGGTTGTAACAGAGCTCTGATGGGTTGTAACTAAGAAACAGCA 180
QY 440 CTTTGATTTGATGATTTTACAGTAAAGCAATTAACCAATCAAGATAGTATTTGTACT 499
DB 181 CTTTGATTTGATGATTTTACAGTAAAGCAATTAACCAATCAAGATAGTATTTGTACT 240
QY 500 TCAGACATGATACATATCTATTTTATAGTATTTAGTGGGTGCTAAATAGCTATTA 559
DB 241 TCAGACATGATACATATCTATTTTATAGTATTTAGTGGGTGCTAAATAGCTATTA 300
QY 560 AATAAGAGAGGAAAAACATGATTAAGTTCATTTAAGTCAACCGTTTGTAGTAAG 619
DB 301 AATAAGAGAGGAAAAACATGATTAAGTTCATTTAAGTCAACCGTTTGTAGTAAG 360
QY 620 AAATCAATTTATCTCCAAACGATTAACGAGTTTACTGAATATCTCAAGTATTGA 679
DB 361 AAATCAATTTATCTCCAAACGATTAACGAGTTTACTGAATATCTCAAGTATTGA 420
QY 680 GACTGTAGTAAATTAAGTTTTTTTGAACAGTACTACTAGCTAATCTTAAGTCTA 739
DB 421 GACTGTAGTAAATTAAGTTTTTTTGAACAGTACTACTAGCTAATCTTAAGTCTA 480
QY 740 TGGATTTGACAGAAATTAATGCTGCT 767
DB 481 TGGATTTGACAGAAATTAATGCTGCT 508

RESULT 14

LACNISIN

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

FEATURES

source

gene

CDS

repeat_unit

CDS

1. 3289
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"
/cissue_11b="ATCC 11454"
1. 174
/gene="nicin"
1. 174
/gene="nicin"
/codon_start=1
/transl_table=11
/protein_id="AAAT7038.1"
/db_xref="GI:149448"
/translation="MSTKDFNLDIVSVSKDSGSPRITSISLCTPGCKTALMGCM
KATNCSTHVS K"
222. 266
/rpt_type=inverted
282. 2840
/codon_start=1
/transl_table=11
/label="ORF A"
/protein_id="AAAT7039.1"
/db_xref="GI:149449"

/translation="MIKSSFKAPLVLNLTISPNDKSFTEYTOVIEVSKNRYE
ULLANPXYDMQKYNAGLKKRKKVLFESIKYKRSYLSRPFGLFSTIGF
SKSQQKMGKTGKIRLDOTWILRLVHMEVDFKSLSPFNANRYKGFQVYT
INSSLEELNRYKYNVOIISFPCNDYOKEDICETVTLQCYGDEYRELSEYGLI
VNHVYLSLQKIDLSFPMNTFLTVBEAIDDKKTIITPLKQKRIQYSEIEBGI
EKUKIYQEMSOILENDNTIQDLSISEINFDVQKQLEHAEFLNATKSVRTY
LDVYDKTIERKGVQVQITELFPISTFGIGAPVYNHPRNDVYSEBETLYSEER
EKYLSMVEAVKNNVINDLDLESHYQKMLEKSELQGLFLMLAKEYKDIILG
DIVGNNGGASGRFSALSPELTSYHRTIVDEBENNEKEITSCIEVLEPNINHAN
VMHTSIMRKVLPFTSTSHNEVLLTNVIGIDEEKERYARDISTOEVLKPFITSMN
KTLFSENEPLVLEISLDKFGULPHELYRBDYIPLVYDEIVISPAKMIKMDVN
SKMTIRELIQSKKEIFIVGDNKRVLSQENPLDMEILSEALVSKSRKDFELQ
YFEDENIINKKEIRVADVVPFIRTRALNGEGRAFIREKRVSRREKLPENMLYL
KLEYSINQNEFLSLYLPDIQKIVANLGNLFELYTDPKPHIRLRKSCDLFLAAYS
ILEILKSRKRNIMSTPDISYDQEVERGCFDLESLAIVPLIKLFOICLH"
repeat_unit
3042..3093
/rpt_type=inverted
repeat_unit
3094..3106
/rpt_type=inverted
-35_signal
3143..3148
-10_signal
3165..3170
-35_signal
3185..3190
-10_signal
3210..3215
3269..3289
/codon_start=1
/label=ORF B
/label_table=11
/protein_id="AA073040.1"
/db_xref="GI:149450"
/translation="MDEVEKF"
BASE COUNT 1263 a 383 c 590 g 1053 t
ORIGIN

Query Match 61.3%; Score 470; DB 1; Length 3289;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

298 ATGAGTACAAAAGATTGTTTAACTGATCGATCGTTGTTCAAGAAAGATTCAGGTGCA 357
1 ATGAGTACAAAAGATTGTTTAACTGATCGATCGTTGTTCAAGAAAGATTCAGGTGCA 60
358 TCACGACGATTAACAAGATTCGCTATGTACACCGGTTGTAAACAGAGACTCTGATG 417
61 TCACGACGATTAACAAGATTCGCTATGTACACCGGTTGTAAACAGAGACTCTGATG 120
418 GGTGTTAATATAAAGCAAGCACTTGCTATGTATTAACGTAAGCAATTAACCAAT 477
121 GGTGTTAATATAAAGCAAGCACTTGCTATGTATTAACGTAAGCAATTAACCAAT 180
478 CAAGAGATGATTTTGTAGTTCAGACATGATCTAATCCATTTTATATAGTATTTA 537
181 CAAGAGATGATTTTGTAGTTCAGACATGATCTAATCCATTTTATATAGTATTTA 240
538 GGGTTCATAAAGCTTATAAATAAAGAGAGAAAAAACAATGATAAAGTTCATTTA 597
241 GGGTTCATAAAGCTTATAAATAAAGAGAGAAAAAACAATGATAAAGTTCATTTA 300
598 AAGCTCAACCGTTTTAGTAAAGAAATCAATTTTATCTCCAAACGATTAACGAGTTT 657
301 AAGCTCAACCGTTTTAGTAAAGAAATCAATTTTATCTCCAAACGATTAACGAGTTT 360
658 CTGATATATCAAGCTGATGAGCTGATGATGATGATGATGATGATGATGATGATG 717
361 CTGATATATCAAGCTGATGAGCTGATGATGATGATGATGATGATGATGATGATG 420
718 TACTAGTAACTCTAACTCTATGATGATGATGATGATGATGATGATGATGATG 767
421 TACTAGTAACTCTAACTCTATGATGATGATGATGATGATGATGATGATGATG 470

RESULT 15
LLINISZ
LOCUS

LLINISZ 2778 bp DNA linear BCT 27-OCT-1999

DEFINITION Lactococcus lactis nisz gene and 3 ORF's.
ACCESSION Y13384
VERSION Y13384.1 GI:3157416
KEYWORDS Nisin Z; Nisz gene.
SOURCE Lactococcus lactis subsp. lactis
ORGANISM Lactococcus lactis subsp. lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1
Immonen, T., Wahlstrom, G., Takala, T. and Saris, P.E.
Evidence for a mosaic structure of the mts481 in Lactococcus lactis
N8
JOURNAL DNA Seq. 9 (5-6), 245-261 (1998)
MEDLINE 99452384
PUBMED 10524753
REFERENCE 2 (bases 1 to 2778)
Immonen, T.
Direct Submission
Submitted (27-MAY-1997) T. Immonen, Institute Of Biotechnology,
Biocenter 1, Viikinkaari 9, P.O.Box 56, 00014 University Of
Helsinki, Finland
FEATURES
source
CDS
1..2778
/organism="Lactococcus lactis subsp. lactis"
/mol_type="genomic DNA"
/strain="N8"
/db_xref="taxon:1360"
/clone="pLEB506"
complement(1..423)
/codon_start=1
/translation="homologous to branched chain amino acid
transporters of LIV-II class"
/protein_id="CA073811.1"
/db_xref="GI:3157417"
/db_xref="SPTRMBL:069437"
/translation="MEKLAGKDYLFISMLFGHFFGAGNLIIPPIHMOEAGALISQA
NEGFLITAVGPPFGIATLGISQNGVELASRYRIVAYITLILVYVIGFPALPR
LATSPFSGISPPSHRIQAPLALFSLFPGTAMFLSR"
complement(660..977)
/codon_start=1
/translation="yjdJ-like protein"
/protein_id="CA073812.1"
/db_xref="GI:3157418"
/db_xref="SPTRMBL:069438"
/translation="MPFWSTKIFGEGKLMNIIRERMLFELSDKGEIIGEMAYVPMN
NSIITHTGVSLDYRGQGLAKKVLVLAGIQRARQOLKLGATCPYAVKFRREKELTD
VLK"
complement(932..1204)
/codon_start=1
/translation="yjdJ-like protein"
/protein_id="CA073813.1"
/db_xref="GI:3157419"
/db_xref="SPTRMBL:069439"
/translation="MDNSKLEGEANERQQLKNDYRKCGEMDIYVVAICENAGEC
VRGNPLVEVSRRKWIIPDNGDVASNOSVINRCSGALKYLAKERN"
2605..2778
/gene="nisZ"
2605..2778
/gene="nisZ"
/gene="nisZ"
/codon_start=1
/translation="nisin Z"
/product="nisin Z"
/protein_id="CA073814.1"
/db_xref="GI:3157420"
/translation="MSTKDFPLDLVSVSKDGSAPRTISLCTPGCKTALMGCMN
KTATNCSSIHVSK"
BASE COUNT 1020 a 399 c 436 g 923 t
ORIGIN

Query Match 59.0%; Score 452.6; DB 1; Length 2778;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 11:19:31 ; Search time 274 Seconds
(without alignments)
7556.456 Million cell updates/sec

Title: US-10-082-618-8

Perfect score: 767
Sequence: 1 agtcgacgaatcttaataa.....tcgacgaatataatgcgcgt 767

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N Geneseq 19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	675	88.0	7454	17	AAT29660
2	643	83.8	7423	20	AAK87792
3	518	67.5	518	11	AAO03354
4	448	58.4	448	14	AAQ50288
5	410.4	53.5	1446	17	AAT29611
6	410.4	53.5	1446	24	ABK14453
7	358.4	46.7	360	13	AAO30070
8	358.4	46.7	360	20	AAK87793

9	319.4	41.6	321	14	AAQ49150
10	319.4	41.6	321	14	AAQ34782
11	319.4	41.6	321	14	AAQ64818
12	202.4	26.4	212	24	AAT69228
13	84.8	11.1	546	14	AAQ50289
14	81.8	10.7	144	18	AAK87829
15	80	10.4	878	24	ABO70186
16	77.8	10.1	930	24	ABO69422
17	77.8	10.1	960	24	ABO67737
18	77.8	10.1	1153020	24	ABO677197
19	77.8	10.1	3011208	24	ABO69245
20	58.4	7.4	2365589	24	ABA90521
21	56.8	7.4	1833	15	AAQ73216
22	54.6	7.1	8056	25	ABZ10246
23	53.4	7.0	11805	24	ABL33748
24	52.8	6.9	100	14	AAQ50292
25	52.2	6.8	50000	24	ABL56644
26	51.4	6.7	8056	25	ABZ10100
27	50.4	6.6	830	11	AAQ03353
28	48.8	6.4	6361	24	ABL33140
29	48.6	6.3	7143	21	AAA70250
30	48.4	6.3	13125	24	ABL70284
31	48.4	6.3	13125	24	ABL33227
32	48.4	6.3	13125	24	ABL34557
33	48.2	6.3	4985	24	ABO75107
34	48.2	6.3	83391	24	ABO67094
35	48	6.3	6065	24	ABL70579
36	48	6.3	6065	24	ABZ10260
37	48	6.3	6065	24	ABK31356
38	48	6.3	11907	24	ABK31320
39	48	6.3	17594	24	ABL34026
40	47.2	6.2	7351	24	ABL32029
41	47	6.1	5452	24	ABL33149
42	46.8	6.1	5303	24	ABL2871
43	46.8	6.1	8056	25	ABZ10246
44	46.8	6.1	17294	24	ABL2986
45	46.6	6.1	7403	22	AA546803

ALIGNMENTS

RESULT 1	AAT29660
ID	AAT29660 standard; DNA; 7454 BP.
XX	AAK87792;
AC	AAT29660;
XX	
DT	01-AUG-1996 (first entry)
XX	
DE	Nisin nisABTClP gene cluster.
XX	
XX	Nisin A; nisa gene; antimicrobial; preservative; antibiotic;
KW	lactibiotic; protein engineering; ds.
XX	
OS	Lactococcus lactis strain NIZO R5.
XX	
Key	Location/Qualifiers
FT	repeat_unit
FT	91..101
FT	/*tag= a
FT	/rpt_type= INVERTED
FT	108..118
FT	/*tag= b
FT	/rpt_type= INVERTED
FT	126..131
FT	/*tag= c
FT	/label= pl-35
FT	152..157
FT	/*tag= d
FT	/label= pl-10
FT	191..197
FT	/*tag= e
FT	206..379
FT	CDS

Bacteriocin gene.
Bacteriocin IL-2 p
Lactococcal bacter
Plasmid pMT1910B D
Lactose operon pro
DNA encoding subcl
Listeria monocytog
Listeria innocua D
Listeria innocua D
Listeria innocua C
Listeria innocua D
Genomic sequence o
Int gene, int gene
Haemolysate of cel
Haemolysate of cel
Human immune syste
Insertion sequence
AmpV genome fragm
Haematopoietic cel
Nucleotide fragmen
Human immune syste
Plasmodium falcipa
Chemically treated
Human immune syste
Human metastasis a
Anopheles gambiae
Human angioendosis
Chemically treated
Human gene regulat
Signal transductio
Human immune syste
Human immune syste
Human immune syste
Haematopoietic cel
Human immune syste
Tumour suppressor

```

FT FT /*tag= f
FT FT /label= nisa
FT FT Sig_peptide 206..274
FT FT /*tag= g
FT FT Mat_peptide 275..376
FT FT /*tag= h
FT FT repeat_unit 127..146
FT FT /*tag= i
FT FT repeat_unit /rpt_type= INVERTED
FT FT 151..171
FT FT /*tag= j
FT FT /rpt_type= INVERTED
FT FT 172..179
FT FT /*tag= k
FT FT CDS 487..3468
FT FT /*tag= l
FT FT RBS 1462..3470
FT FT /*tag= m
FT FT CDS 3479..5281
FT FT /*tag= n
FT FT /label= nisp
FT FT 5260..5267
FT FT RBS
FT FT CDS 5274..6518
FT FT /*tag= o
FT FT CDS 6497..6505
FT FT /*tag= q
FT FT CDS 6515..7252
FT FT /*tag= r
FT FT RBS 7237..7245
FT FT /*tag= s
FT FT CDS 7254..7454
FT FT /*tag= t
FT FT /label= nisp
FT FT /note= "3' region of nisp is incomplete"
PN W09616180-A1.
XX 30-MAY-1996.
XX 20-NOV-1995; 95WO-GB02699.
XX 19-NOV-1994; 94GB-0023404.
XX (BIOT-) BIOTECHNOLOGY & BIOLOGICAL SCI RES COUNC.
XX PA
XX PI Dodd HM, Gasson MJ;
XX XX
XX WPI, 1996-268616/27.
XX P-PSDB, AAR95267, AAR95268, AAR95269, AAR95270, AAR95271.
XX
XX Making cell which expresses nisin but does not contain natural nisa
XX PT gene - by providing cell with variant nisa gene, and genes for nisin
XX PT modification, secretion and immunity
XX PS
XX Disclosure; Fig 7; 69pp; English.
XX
XX The gene cluster nisaBTPRK (see AAT9660 and AAT9661) of Lactococcus
XX CC lactis includes the nisa gene coding for pre-nisin A (AAR95267, see
XX CC also AAR95263) and the genes for nisin modification, secretion and
XX CC immunity. nisaB (AAR95268) and nisc (AAR95270) are believed to be
XX CC involved in reactions that modify pre-nisin; nisp (AAR95269) is
XX CC similar to a transport ATPase and is involved in translocation of
XX CC nisin out of the cell; nisi (AAR95271) is involved in immunity to
XX CC nisin. Replacement of the natural, chromosomal copy of the nisa
XX CC gene with a variant nisa gene allows produ. of high levels of nisin
XX CC A variants in Lactococcus lactis hosts.
XX
XX Sequence 7454 BP; 2741 A; 920 C; 1376 G; 2417 T; 0 other;
SQ

```

```

Query Match 88.0%; Score 675; DB 17; Length 7454;
Best Local Similarity 100.0%; Pred. No. 1,4e-126;
Matches 675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 AATGACCTAGTCTTATTAATCTACTGACATATGAAAACATTAACAATCTAAACAGCTTT 152
DB 1 AATGACCTAGTCTTATTAATCTACTGACATATGAAAACATTAACAATCTAAACAGCTTT 60
OY 153 AATCTATCTGAGAAAGATTTGGTAATATATTTTTCATTAACGGAGATTAATAA 212
DB 61 AATCTATCTGAGAAAGATTTGGTAATATATTTTTCATTAACGGAGATTAATAA 120
OY 213 CGGCTCTGATTTAAATTTCTGAAGTTTGTAGATACATATGTTTCGTCAGAGAACTACAA 272
DB 121 CGGCTCTGATTTAAATTTCTGAAGTTTGTAGATACATATGTTTCGTCAGAGAACTACAA 180
OY 273 AATAAATTATTAAGAGGACCTCAAAATAGTACAAAGATTTTAACTTGATTTGATTC 332
DB 181 AATAAATTATTAAGAGGACCTCAAAATAGTACAAAGATTTTAACTTGATTTGATTC 240
OY 333 TGTTCGAGAAAGATTCAGGTGATCCACGACCTTCAAGTATTTGGCTATGTACACC 392
DB 241 TGTTCGAGAAAGATTCAGGTGATCCACGACCTTCAAGTATTTGGCTATGTACACC 300
OY 393 CGTTGTAAAAACAGAGAGCTGTATGGGTGTATACATGAAAACAGCACTTGCTATGTAG 452
DB 301 CGTTGTAAAAACAGAGAGCTGTATGGGTGTATACATGAAAACAGCACTTGCTATGTAG 360
OY 453 TATTCAGGTAAAGCAATTAACCAATCAAAAGATGATTTTGTATGTTCAAGATGATA 512
DB 361 TATTCAGGTAAAGCAATTAACCAATCAAAAGATGATTTTGTATGTTCAAGATGATA 420
OY 513 CTATCCTATTTTATTAAGTTATTTAGGTTCTAAATGCTTATTAATAAATTAAGAGAGA 572
DB 421 CTATCCTATTTTATTAAGTTATTTAGGTTCTAAATGCTTATTAATAAATTAAGAGAGA 480
OY 573 AAAACATGATTAATAAGTTCAATTTAAAGCTCAACGTTTATTAAGTAAGAAATCAATTTTA 632
DB 481 AAAACATGATTAATAAGTTCAATTTAAAGCTCAACGTTTATTAAGTAAGAAATCAATTTTA 540
OY 633 TCTCCAAACGATTAACGAGTTTACTGAATATTAAGTCAAGTCAATGAGCTGTAAATGTA 692
DB 541 TCTCCAAACGATTAACGAGTTTACTGAATATTAAGTCAAGTCAATGAGCTGTAAATGTA 600
OY 693 AATAAAGTTTATTTGGAACAGTTACTACTAGCTAATCCTAAACCTAAGACTTATGACG 752
DB 601 AATAAAGTTTATTTGGAACAGTTACTACTAGCTAATCCTAAACCTAAGACTTATGACG 660
OY 753 AATATATATGCTGT 767
DB 661 AATATATATGCTGT 675

RESULT 2
AAX87792
ID AAX87792 standard; DNA; 7423 BP.
XX AAX87792;
XX
XX 09-NOV-1999 (first entry)
XX
XX Nisin gene region of Lactobacillus lactis.
XX DE
XX Nisin A; nls A; nisin B; nls B; nisin T; nls T; nisin C; nls C;
XX KW nisin I; nls I; lantibiotic; animal performance; growth rate;
XX feed conversion; bacteriocin; probiotic; ss.
XX
XX Lactobacillus lactis.
XX
XX Key Location/Qualifiers
XX RBS 159..165
XX FT /*tag= a
XX CDS 174..347

```


FT FT /tag= b
FT /note= "nis A; encodes AA06665"
FT 440..441
RBS /tag= c
FT 444..447
RBS /tag= d
FT 455..3436
CDS /tag= e
FT /note= "nis B; encodes AA06666"
FT 3430..3433
RBS /tag= f
FT 3435..3438
RBS /tag= g
FT 3447..5249
CDS /tag= h
FT /note= "nis T; encodes AA06667"
FT 5228
RBS /tag= i
FT 5242..6486
CDS /tag= j
FT /note= "nis C; encodes AA06668"
FT 6483..7220
FT /tag= k
FT /note= "nis I; encodes AA06669"

XX MO9941978-A1.
PN 26-AUG-1999.
PD 12-FEB-1999; 99WO-IB00250.
PF 18-FEB-1998; 96GB-0003424.
PR (PIIZ) PFIZER INC.
PA (PIIZ) PFIZER LTD.
XX
XX Flanagan AJ, Haxell MA, Rolph TP;
PI WPI, 1999-527402/44.
DR P-PBDB; AA06665, AA06666, AA06667, AA06668, AA06669.
XX Novel performance enhancing method resulting in improved growth
PT rates and feed conversion efficiencies in animals
PT
XX
XX Disclosure; Page 55-57; 79pp; English.

XX This is the nisn gene region of Lactobacillus lactis NIZO R5,
CC including the nis A, nis B, nis T, nic C and nis I genes that
CC respectively code for nistins A, B, T, C and I (see AA06665-69).
CC The invention relates to methods of enhancing performance in an
CC animal by administering a bacterium capable of expressing a
CC performance enhancing polypeptide such as nisn A or nisn Z. The
CC enhanced performance results in improved growth rates and feed
CC conversion efficiencies. The bacterium, which may be genetically
CC modified to express the performance enhancing polypeptide, is
CC administered to the gastrointestinal tract, especially to an
CC embryo or neonatal animal. If the polypeptide is nisn, it may
CC also inhibit ruminal methane, decrease acetate to propionate ratios
CC and prevent amino acid deamination. Administration of an
CC appropriate bacterium which may act as a probiotic may also help
CC control enteric pathogens in poultry. The polypeptide is produced
CC continuously in the gut, maintaining a constant level. The protein
CC is eventually degraded, leaving no residues in the meat.
XX

Sequence 7423 BP; 2729 A; 913 C; 1374 G; 2407 T; 0 other;

Query Match	83.8%	Score 643	DB 20	Length 7423
Best Local Similarity	100.0%	Pred. No.	3.8e-120	
Matches	643	Conservative	0	Mismatches 0; Indels 0; Gaps 0

QY	125	GAACATTACCAATCTTAAAAAGCTCTTATTTCTATCTTGAGAAAGTATTTGGTATATATA	184
Db	1	GAACATTACCAATCTTAAAAAGCTCTTATTTCTATCTTGAGAAAGTATTTGGTATATATA	60

OY	185	TTATTTGCGATAACGGGACGACATATAAAGCGCTCGATTAATCTGAAGTTGTGAT	244
Db	61	TTATTTGCGATTAACGGGACGACATATAAAGCGCTCGATTAATCTGAAGTTGTGAT	120
OY	245	ACAATGATTTCTGTCGAAGAACTACAAATTAATTAAGAGGCACTCAAAATGAATA	304
Db	121	ACAATGATTTCTGTCGAAGAACTACAAATTAATTAAGAGGCACTCAAAATGAATA	180
OY	305	CAAAAATTTTAACCTTGGATTTGGTATCTGTTTGGAGAAAAGATTCAGTGCATCACCAC	364
Db	181	CAAAAATTTTAACCTTGGATTTGGTATCTGTTTGGAGAAAAGATTCAGTGCATCACCAC	240
OY	365	GCATTCAGATGATTTGCGATATGACACCCGGTGTAAAAACAGAGCTCTGATGGGTGTA	424
Db	241	GCATTCAGATGATTTGCGATATGACACCCGGTGTAAAAACAGAGCTCTGATGGGTGTA	300
OY	425	ACATGAAAAACAGAACTGTCTATTGTAGTATTCAGTAGCAATTAACCAATCAAGAA	484
Db	301	ACATGAAAAACAGAACTGTCTATTGTAGTATTCAGTAGCAATTAACCAATCAAGAA	360
OY	485	TAGTATTTTGTAGTTGACAGCATGATCTATCCATTTTATTAAGTATTTAGGGTTGC	544
Db	361	TAGTATTTTGTAGTTGACAGCATGATCTATCCATTTTATTAAGTATTTAGGGTTGC	420
OY	545	TAAATAGCTTATPAAAAATPAAAGAGAGGAAAAACATGATAAAAGTCAATTAAGCTCA	604
Db	421	TAAATAGCTTATPAAAAATPAAAGAGAGGAAAAACATGATAAAAGTCAATTAAGCTCA	480
OY	605	ACCGTTTTTGTAGTAAGAAATCAATTTTATCTCCAACGATPAAACGAGTTTTACTGATA	664
Db	481	ACCGTTTTTGTAGTAAGAAATCAATTTTATCTCCAACGATPAAACGAGTTTTACTGATA	540
OY	665	TACTCAAGTCATTTGAGACTGTAAAGTAAAAATPAAAGTTTTTTGGACAAGTTACTACAGC	724
Db	541	TACTCAAGTCATTTGAGACTGTAAAGTAAAAATPAAAGTTTTTTGGACAAGTTACTACAGC	600
OY	725	TAAATCTAAACTATGATGTATGATATGCAAAATPATAATGCTGGT	767
Db	601	TAAATCTAAACTATGATGTATGATATGCAAAATPATAATGCTGGT	643

XX	RESULT 3
AAQ03354	ID
AAQ03354 standard; DNA; 518 BP.	
XX	
AC	AAQ03354;
XX	
DT	04-AUG-1990 (first entry)
XX	
DE	Nucleotide fragment containing gene encoding nisin and its precursor
DE	peptide derived from Streptococcus lactis ATCC 11454.
XX	
KW	Nisin precursor peptide; peptide leader fragment;
KV	peptide leader sequence; post-translational modification;
XX	Streptococcus lactis ATCC 11454.
OS	Streptococcus lactis.
XX	
FH	Key
FH	Location/qualifiers
FT	CDS
FT	298..471
FT	/+tag= a
FT	298..367
FT	/+tag= b
FT	/note="leader fragment responsible for inducing post-
FT	translational modification"
RBS	283..289
FT	/+tag= c
FT	260..297
FT	/+tag= d
FT	472..518
FT	/+tag= e
misc_feature	(416..435)

```
/*tag= f
/note="20mer probe"
```

MO9000558-A

25-JAN-1990

30-JUN-1989; 89WO-US02820

05-JUL-1988; 88US-0214959

(UYMA-) UNIV OF MARYLAND.

Hansen NJ;

WPI; 1990-051685/07

1
2
3
4
5
6
7
8
9
10
11
12

including post-tr

Disclosure; Fig 3; 19pp; English.

It contains a gene leader fragment encoding peptide leader sequence which

Thr. The nisin gene was found in S. Lactis using 20mer probe (see Fig. 4). A DNA fragment is expected encoding a polypeptide

attached to a leader fragment, this is then inserted into an expression

proteins containing dehydroalanine and dehydrobutyrine which can now be bound to specific substrates.

designed. Tag b is claimed in the patent in claim 6.

Sequence 518 BP; 187 A; 74 C; 89 G; 168 T; 0 other;

every Match	67.5%;	Score 518;	DB 11;	Length 518;
-------------	--------	------------	--------	-------------

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGTTGACGATAATTAAATAATTATCTGATTTCTAGTTCCTGAATAATA 60

1 AGTTGACGATATTTAATAATTTTATTAAATATCTGATTTTCTAGTTCCCTGATAATATA 60

61 GAGATAGGTTTATTTGAGTCTTAGACATACTTGATGACCTAGTCTTATAACTATACTGAC 12

61 GAGATAGGTTTATTGAGCTTTAGACATACTTGAATGACCTAGTCTTATACTATACTGAC 12

121 AATAGAAACATTACCAATCTAAACAGTCTTAATTCTATCTTGAGAAAGTATGCTAAT 18

121 AATAGAACATTAAACAATCTAAACAGTCTTAATTCTATCTTGAGAAAGTATTGGTAAT 18

181 AATAATATGTCGATAACGGAGCATATAAAGGCTCTGATTAAATCTGAAGTTTCTT 24

181 AATAATTATGTGATTAACGGGACATTAATAAACGGCTCTGATTAAATTCTGAAGTTTGT 24

241 AGATACCAATGATTTTCGTTGAGGGACTACAAATAAATTATAAGGAGGCAC TCAAATG 300

241 AGATACATGATTTCGTCGAAGGAACACAAATAATTATAGGAGGCATCAAATG 300

301 AGTACCAAGATTTTAACTTGGATTTCGTATCTGTTTTCGAAGAAGATTCAGGTCATCA 36

301 AGTACAAAGAATTAACTTGGATTGGTATCTGTTTCGAGAAAGATTCAGGTGCATCA 36

361 CCACGCATTACAGTATTTCCGCTATGTACACCCGGTGTGAAACAGGAGCTCTGATGGGT 42

361 CCACGATTACAAGTAATTTCGCTAAGTACACCCCGTTGTAAACAGGAGCTCTGATGGGT 42

421 TGTACCATGAAACAGCAACTTGTCACTGTGAGTATTCACGTAGCAAAATACCAATCAA 48

421 TGTACATGAACAACAGCACTTGTCACTGTAGTATTACCGTAAGCAATTAACCAATCAA 48

481 AGGATAGTATTTGTAGTTCAGACATGGATACTATCC 518

```

Db      481 AGGATAGTATTGTGTACAGCAGCATGACTATCC 518
RESULT 4
AAQ50288
ID      AAQ50288 standard; DNA; 448 BP.
XX
AC      AAQ50288;
DT      25-MAR-2003 (updated)
DT      03-MAY-1994 (first entry)
XX
DE      nlsA gene and flanking regions.
KM      Nlsin; antibiotic; food; spoilage; pathogens; industrial processes;
KW      lantibiotic; Lactococcus lactis; ss.
OS      Lactococcus lactis.
XX
FH      Key          Location/Qualifiers
FT      RBS          73..79
FT                      /*tag= a
FT      CDS           88..259
FT                      /*tag= b
FT      RBS          356..361
FT                      /*tag= c
FT      CDS           369..448
FT                      /*tag= d
FT                      /label= nlsB gene.
FT                      /note= "continues to end of sequence. The gene is
FT                        repeat_region
FT                        309..353
FT                        /*tag= e
FT                        /note= "Inverted repeat."
XX
PN      W09320213-A1.
PD      14-OCT-1993.
PF      01-APR-1993;   93WO-GH00676.
PR      02-APR-1992;   92GB-0007267.
XX
PA      (AGRI-) AGRIC & FOOD RES COUNCIL.
PI      Dodd HM, Gaeson WJ;
PS      WPI; 1993-336920/42.
XX
PT      New organisms able to express genes for nisin maturation, but not
PT      native pre-nisin - useful for producing variant nisin(s) for
PT      antimicrobial use in food industry
XX
PS      Disclosure; Figure 2; 62bp; English.
XX
CC      Nisin is a highly modified peptide antibiotic produced by certain
CC      bacteria e.g. strains of Lactococcus lactis. It has efficient
CC      antimicrobial activity against a wide range of Gram positive
CC      bacteria which are food pathogens or spoilage organisms. The genes
CC      involved in nisin maturation can be used to transform other microbes
CC      which are unable to secrete the natural nlsA nisin. These organisms
CC      can however modify nisin and translocate it out of the cell. This
CC      provides an effective means of producing variant nisms which are
CC      useful as antimicrobial agents in conditions of high pH where
CC      natural nisin is ring-opened and loses its antimicrobial properties.
CC      (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ      Sequence 448 BP; 167 A; 64 C; 80 G; 137 T; 0 other;
Query Match          58.4%; Score 448; DB 14; Length 448;
Best Local Similarity 100.0%; Pred. NO. 4,4e-81;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 211 AACGCTCTGATTAATCTGAAGTTGTAGATCAATGATTTGTTGCGAAGAACTAC 270
 DB 1 AACGCTCTGATTAATCTGAAGTTGTAGATCAATGATTTGTTGCGAAGAACTAC 60
 OY 271 AAAATAAATTAAGAGGCACTCAAAATGAGTACAAAAGATTTTAACTTGATTTGTA 330
 DB 61 AAAATAAATTAAGAGGCACTCAAAATGAGTACAAAAGATTTTAACTTGATTTGTA 120
 OY 331 TCTGTTTGAAGAAGATTCAGTGCATCAACAGCATTACAGATTTTCGTAATGACA 390
 DB 121 TCTGTTTGAAGAAGATTCAGTGCATCAACAGCATTACAGATTTTCGTAATGACA 180
 OY 391 CCGGTTTGAAGAAGATTCAGTGCATCAACAGCATTACAGATTTTCGTAATGACA 450
 DB 181 CCGGTTTGAAGAAGATTCAGTGCATCAACAGCATTACAGATTTTCGTAATGACA 240
 OY 451 AGTATTCAGTAAAGCAATTAACCAATCAAGATGATTTTGTAGTTGAGACATGGA 510
 DB 241 AGTATTCAGTAAAGCAATTAACCAATCAAGATGATTTTGTAGTTGAGACATGGA 300
 OY 511 TACTATCTATTTTATTAAGTATTTAGGCTTGAATAGCTTTTAAATAAAGAGAG 570
 DB 301 TACTATCTATTTTATTAAGTATTTAGGCTTGAATAGCTTTTAAATAAAGAGAG 360
 OY 571 GAAAAAAGATGATTAAGTTGATTTAAAGCTCAACCGTTTAAAGAAATACATTT 630
 DB 361 GAAAAAAGATGATTAAGTTGATTTAAAGCTCAACCGTTTAAAGAAATACATTT 420
 OY 631 TATCTCAACGATTAACGAGATTTTAC 658
 DB 421 TATCTCAACGATTAACGAGATTTTAC 448

RESULT 5
 AAT29611
 ID AAT29611 standard; DNA; 1446 BP.
 AC AAT29611;

XX 08-JAN-1997 (first entry)
 DT
 XX
 DE Nisin A promoter fragment.

XX Nisin Z; lactococcus; lactic acid bacterium; promoter; membrane protein;
 KM antimicrobial peptide; nisin A; cell lysis; enzyme; fermented foodstuff;
 KW extracellular protein; dairy product; cheese; yoghurt; inducer; ss.

XX Lactococcus lactis.

OS
 XX
 FH Key Location/Qualifiers
 FT -35_signal 1252..1257
 FT -10_signal 1278..1283
 FT /*tag= b
 FT 1332..1446
 FT /*tag= C
 FT /product= Nisin A fragment

EP712935-AZ.

PD 22-MAY-1996.

PF 17-NOV-1995; 95EP-0203153.

PR 18-NOV-1994; 94NL-0001935.

PR 18-NOV-1994; 94NL-0001934.

PA (NIZO-) NIZO STICHTING NEDERLANDS INST ZUIVELOND.

XX De Vos WM, Kuipers OP;

XX MPI; 1996-252638/26.

XX Controlled lysis of transformed lactic acid bacteria to release
 PT intracellular proteins or RNA - by addn. of inducer to activate
 PT promoter to express antimicrobial peptide, useful for prodn. of
 PT dairy prods.

Example 2; Fig 4b; 29pp; English.

CC This sequence represents the nisin A promoter isolated from *Lactococcus*
 CC lactis. This sequence is used to control the DNA fragment of the method
 CC of the invention. The method of the invention is for the controlled
 CC expression of a DNA fragment containing one or more genes of desired
 CC characteristics, in a lactic acid bacterium. The DNA fragment is under
 CC the control of a promoter for a microbial gene that encodes an
 CC antimicrobial peptide. In the method the gene on the DNA fragment are
 CC expressed by the addition of a suitable inducer for transcription
 CC activation, such as nisin A, or nisin Z (see AAR95245). The inducer
 CC used is an antimicrobial peptide produced by lactic acid bacterium. The
 CC expression of the DNA fragment causes the lysis of the microorganism
 CC used, or alternatively results in the release of intracellular proteins
 CC into the medium. The method can be used for the production of proteins
 CC (such as enzymes, membrane proteins, extracellular proteins and
 CC antimicrobial proteins) or RNA. The method is particularly useful for
 CC producing dairy products containing a desired protein. The release of
 CC the intracellular enzymes provides accelerated maturing of fermented
 CC foodstuffs (such as cheese or yoghurt). The bacterium containing the
 CC DNA fragment can be used to determine the concentration of a inducer in
 CC a medium. Cell lysis can be induced at any point during culture by
 CC addition of the inducer.

Sequence 1446 BP; 488 A; 233 C; 244 G; 481 T; 0 other;

Query Match 53.5%; Score 410.4; DB 17; Length 1446;
 Best Local Similarity 99.8%; Pred. No. 1.7e-73;
 Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AGTGACGAATATTAATTAATTTTATTAATATCTGATTTTCTAGTTCGTAATATATA 60
 DB 1035 AGTGACGAATATTAATTAATTTTATTAATATCTGATTTTCTAGTTCGTAATATATA 1094
 OY 61 GAGATAGGTTTATTTAGTCTTGAACATACCTGTAATGACCTTGAATCTATATCTGAC 120
 DB 1095 GAGATAGGTTTATTTAGTCTTGAACATACCTGTAATGACCTTGAATCTATATCTGAC 1154
 OY 121 AATGAAACATTAACAAATCTTAAACGCTTAAATCTTGAAGAGATTTGTAAT 180
 DB 1155 AATGAAACATTAACAAATCTTAAACGCTTAAATCTTGAAGAGATTTGTAAT 1214
 OY 181 AATATTTATTTGATTAACGAGCATTAATAACGCTCTGATTAATCTGAAGTTTGT 240
 DB 1215 AATATTTATTTGATTAACGAGCATTAATAACGCTCTGATTAATCTGAAGTTTGT 1274
 OY 241 AGATACAAATGATTTGTTGTAAGGAACCTACAAATTAATTAAGAGGCACTCAAAATG 300
 DB 1275 AGATACAAATGATTTGTTGTAAGGAACCTACAAATTAATTAAGAGGCACTCAAAATG 1334
 OY 301 AGTACAAAAGATTTTAATCTTGATTTGTAATCTGTTTGAAGAAAGATTCAGTGCATCA 360
 DB 1335 AGTACAAAAGATTTTAATCTTGATTTGTAATCTGTTTGAAGAAAGATTCAGTGCATCA 1394
 OY 361 CCAGCATTAACAAGATTTGCTATGTACACCGGTTGTAATAACAGAGACTC 412
 DB 1395 CCAGCATTAACAAGATTTGCTATGTACACCGGTTGTAATAACAGAGACTC 1446

RESULT 6

ABK14453 standard; DNA; 1446 BP.

XX ABK14453;

XX 08-MAY-2002 (first entry)

DE Lactococcus lactis nisa promoter sequence, from Tn5276.
 XX
 KW Gene expression control; antimicrobial peptide; food product; nisin A;
 KW nisa; promoter; Tn5276; transposon; lantibiotic; ds.
 XX
 OS Lactococcus lactis.
 FH
 FH Key Location/Qualifiers
 FT -35_signal 1252..1257
 FT /tag= a
 FT /standard_name= "-35 box"
 FT -10_signal 1278..1283
 FT /tag= b
 FT /standard_name= "-10 box"
 FT RBS 1317..1323
 FT /tag= c
 FT /standard_name= "Shine-Dalgarno (SD) sequence"
 XX
 PN EPI162271-A2.
 XX
 PD 12-DEC-2001.
 XX
 PF 17-NOV-1995; 2001EP-0201701.
 XX
 PR 18-NOV-1994; 94NL-0001934.
 PR 18-NOV-1994; 94NL-0001935.
 PR 17-NOV-1995; 95EP-0203153.
 XX
 XX (NIZO-) NIZO STICHTING NEDERLANDS INST ZUIVELOND.
 XX
 PI Kuipers OP, De Vos WM;
 FI WPI; 2002-107870/15.
 DR
 XX
 XX Controlled expression of a gene for production of a protein comprises
 PT a DNA fragment under control of a promoter from an
 PT antimicrobial peptide, where the peptide induces the promoter -
 XX
 PS Example 2; Fig 4b; 31pp; English.
 XX
 CC The present invention relates to a new method for controlling expression
 CC of a gene. The method of the invention involves providing a DNA fragment
 CC comprising a gene under transcriptional control of a promoter, where the
 CC promoter is obtained from a Gram-positive gene for an antimicrobial
 CC peptide or a precursor of the peptide. The gene is found in a gene
 CC cluster for the synthesis of the antimicrobial peptide and the promoter
 CC is inducible by the peptide. The antimicrobial peptide of the invention
 CC is acceptable in food products. The method provides a homologous
 CC expression system in Lactococcus, which is able to give strict/absolute
 CC regulation of the gene expression using (in low concentrations) suitable
 CC inducers. High production of the desired protein, aminopeptidase N in
 CC experiments, gave a specific activity of 25000 nmol/mg/min using 0.5 g/l
 CC nisin A as the inducer, a level that has never been reached before by use
 CC of other (constitutive) promoters. The present nucleic acid sequence
 CC represents the nisa promoter sequence of the invention.
 XX
 SQ Sequence 1446 BP; 488 A; 233 C; 244 G; 481 T; 0 other;

QY 181 AATATTATGTCGATACGCGACATTAATAAGGCTCTGATTAATTCTGAAGTTGTT 240
 DB 1215 AATATTATGTCGATACGCGACATTAATAAGGCTCTGATTAATTCTGAAGTTGTT 1274
 QY 241 AGATACATGATTTGCTTGGAAGAACTACCAAAATTAATTATAGAGGCACTCAAAATG 300
 DB 1275 AGATACATGATTTGCTTGGAAGAACTACCAAAATTAATTATAGAGGCACTCAAAATG 1334
 QY 301 AGTACAAAAGATTTAACTTGATTTGATCTGTTGGAAGAAAGATTACAGTGCATCA 360
 DB 1335 AGTACAAAAGATTTAACTTGATTTGATCTGTTGGAAGAAAGATTACAGTGCATCA 1394
 QY 361 CCAGCCATTACAGATTTGCTATGTACACCCGGTTGTAACAGAGCTC 412
 DB 1395 CCAGCCATTACAGATTTGCTATGTACACCCCGTTGTAACAGAGCTC 1446
 RESULT 7
 AAQ30070
 ID AAQ30070 strand; DNA; 360 BP.
 XX
 AC AAQ30070;
 XX
 DT 25-MAR-2003 (updated)
 DT 02-APR-1993 (first entry)
 XX
 DE Sequence of the nisz gene isolated from L. Lactis NIZO 22186.
 XX
 XX Lantibiotic; nisin Z; nisin A; analogue; food preservative; ss.
 XX
 OS Lactococcus lactis.
 FH
 FH Key Location/Qualifiers
 FT RBS 106..112
 FT /tag= a
 FT CDS 121..189
 FT /tag= b
 FT /product= leader peptide
 FT mat_peptide 190..294
 FT /tag= c
 FT misc_feature 21..40
 FT /tag= d
 FT /label= primer
 FT misc_feature 320..340
 FT /tag= e
 FT /label= primer
 XX
 PA (NIZO-) NIZO STICHTING NEDERLANDS INST ZUIVELOND.
 XX
 PI De Vos WM, Kuipers OP, Siezen RJ;
 FI WPI; 1992-382116/46.
 DR P-PeDB; AAR26298.
 DR
 XX
 PT New lantibiotic cpds. related to nisin A - and Lactococcus
 PT strains which produce them, useful as preservatives for foods and
 PT animal feeds
 PS
 XX Example; Fig 3; 42pp; English.
 CC Total DNA was isolated from the L. lactis strains NIZO 22186 and
 CC NIZO R5. With the aid of the nisa gene of strain NIZO R5 as a probe,
 CC a 4.5 kb HindIII fragment was identified in the total DNA strain
 CC NIZO 22186, which was then cloned in M13 mp18. The DNA sequence of

CC the gene for nisin Z production (nisZ) was determined by making use
 CC of oligos complementary to the 5' and 3' flanking sequences of the
 CC nisa gene. The nucleotide sequence of the nisz gene is found to be
 CC identical to that of the nisa gene with the exception of a C to A
 CC transversion in posn. 148 which results in the replacement of AA
 CC His37 by Asn27. The above indicate that the structure of nisin Z is
 CC as shown in AAR28299.
 CC (Updated on 25-MAR-2003 to correct PN field.)

CC Sequence 360 BP; 129 A; 52 C; 66 G; 113 T; 0 other;

Query Match 46.7%; Score 358.4; DB 13; Length 360;
 Best Local Similarity 99.7%; Pred. No. 4.3e-63;
 Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 178 AATAATATTATTTGTCATACCGGACATTAATTAAGGAGGCACTGCA 237
 DB 1 AATAATATTATTTGTCATACCGGACATTAATTAAGGAGGCACTGCA 60
 QY 238 GTTAGATACATGATTTGTCGAGAGAACTAATAATTAATTAAGGAGGCACTGCA 297
 DB 61 GTTAGATACATGATTTGTCGAGAGAACTAATAATTAATTAAGGAGGCACTGCA 120
 QY 298 ATGAGTACAAAAGATTTTGAATTTGATCTGTTTGAAGAAAGATTCAAGTGCA 357
 DB 121 ATGAGTACAAAAGATTTTGAATTTGATCTGTTTGAAGAAAGATTCAAGTGCA 180
 QY 358 TCACCGCATTAACAAGTATTCGTATGTACACCGGTTTGAAGAAAGAGCTGTGATG 417
 DB 181 TCACCGCATTAACAAGTATTCGTATGTACACCGGTTTGAAGAAAGAGCTGTGATG 240
 QY 418 GGTGTAACATGAAGAAAGCAACTGTGATGTAGTATTCAGTAAGCAATTAACCAAT 477
 DB 241 GGTGTAACATGAAGAAAGCAACTGTGATGTAGTATTCAGTAAGCAATTAACCAAT 300
 QY 478 CAAAGATAGTATTTTGTAGTTCAGACATGATCTATCTATTTTATAAGTATTTA 537
 DB 301 CAAAGATAGTATTTTGTAGTTCAGACATGATCTATCTATTTTATAAGTATTTA 360

RESULT 8
 AAX87793
 ID AAX87793 standard; DNA; 360 BP.

AC AAX87793;
 XX
 DT 09-NOV-1999 (first entry)
 XX
 DE Nisin Z gene of *Lactobacillus lactis*.
 XX
 KM Nisin Z; nis Z; lantibiotic; animal performance; growth rate;
 XX feed conversion; bacteriocin; probiotic; ss.
 OS

Lactobacillus lactis.

Key Location/Qualifiers
 FT CDS 121..294
 FT /tag= a
 FT /note= "nis Z"

MO9941978-A1.

26-AUG-1999.

12-FEB-1999; 99WO-1B00250.

18-FEB-1998; 98GB-0003424.

(PFIZ) PFIZER INC.
 PA (PFIZ) PFIZER LTD.

Planagan AV, Haxell MA, Rolph TP;

DR WPI, 1999-527402/44.

DR P-PSDB; AAY06670.

PT Novel performance enhancing method resulting in improved growth
 PT rates and feed conversion efficiencies in animals

PS Disclosure; Page 71; 79pp; English.

CC This is the nisin Z (nis Z) gene region of *Lactobacillus lactis*
 CC strain 22186. Nisin Z is a natural analogue of nisin A (see
 CC AAY06665), a lantibiotic-containing bacteriocin. The invention
 CC relates to methods of enhancing performance in an animal by
 CC administering a bacterium capable of expressing a performance
 CC enhancing polypeptide such as nisin A or nisin Z. The enhanced
 CC performance results in improved growth rates and feed conversion
 CC efficiencies. The bacterium, which may be genetically modified to
 CC express the performance enhancing polypeptide, is administered to
 CC the gastrointestinal tract, especially to an embryo or neonatal
 CC animal. If the polypeptide is nisin, it may also inhibit ruminal
 CC methane, decrease acetate to propionate ratios and prevent amino
 CC acid deamination. Administration of an appropriate bacterium which
 CC may act as a probiotic may also help control enteric pathogens in
 CC poultry. The polypeptide is produced continuously in the gut,
 CC maintaining a constant level. The protein is eventually degraded,
 CC leaving no residues in the meat.

CC Sequence 360 BP; 129 A; 52 C; 66 G; 113 T; 0 other;

Query Match 46.7%; Score 358.4; DB 20; Length 360;
 Best Local Similarity 99.7%; Pred. No. 4.3e-63;
 Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 178 AATAATATTATTTGTCATACCGGACATTAATTAAGGAGGCACTGCA 237
 DB 1 AATAATATTATTTGTCATACCGGACATTAATTAAGGAGGCACTGCA 60
 QY 238 GTTAGATACATGATTTGTCGAGAGAACTAATAATTAATTAAGGAGGCACTGCA 297
 DB 61 GTTAGATACATGATTTGTCGAGAGAACTAATAATTAATTAAGGAGGCACTGCA 120
 QY 298 ATGAGTACAAAAGATTTTGAATTTGATCTGTTTGAAGAAAGATTCAAGTGCA 357
 DB 121 ATGAGTACAAAAGATTTTGAATTTGATCTGTTTGAAGAAAGATTCAAGTGCA 180
 QY 358 TCACCGCATTAACAAGTATTCGTATGTACACCGGTTTGAAGAAAGAGCTGTGATG 417
 DB 181 TCACCGCATTAACAAGTATTCGTATGTACACCGGTTTGAAGAAAGAGCTGTGATG 240
 QY 418 GGTGTAACATGAAGAAAGCAACTGTGATGTAGTATTCAGTAAGCAATTAACCAAT 477
 DB 241 GGTGTAACATGAAGAAAGCAACTGTGATGTAGTATTCAGTAAGCAATTAACCAAT 300
 QY 478 CAAAGATAGTATTTTGTAGTTCAGACATGATCTATCTATTTTATAAGTATTTA 537
 DB 301 CAAAGATAGTATTTTGTAGTTCAGACATGATCTATCTATTTTATAAGTATTTA 360

RESULT 9

AAQ49150
 ID AAQ49150 standard; DNA; 321 BP.

AAQ49150;

25-MAR-2003 (updated)
 DT 10-MAR-1994 (first entry)

Bacteriocin gene.

Bacteriocin; LL-2; gram positive bacteria; antimicrobial; food;
 KM PCR; polymerase chain reaction; amplification; inhibit; ss.

Lactococcus lactis.

Key Location/Qualifiers
CDS 101..274
/*tag= a
/product= Bacteriocin

US5232849-A.
03-AUG-1993.
14-MAY-1992; 92US-0882079.
01-JUL-1991; 91US-0721774.
14-MAY-1992; 92US-0882079.
(UNIL) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
Henderson JT, Marugg JD, Van MASSENAR PD, Vedamuthu ER;
WPI, 1993-287077/36.
P-PSDB; AAR41280.

Bacteriocin from *Lactococcus lactis* subspecies *lactis* - useful as
inhibitory against Gram-positive bacteria

Claim 1; Page 13-14 (col 15,16,17,18); 14pp; English.

The sequence (AAQ49150) was amplified using primers (AAQ49151-52). The
encoded protein inhibits selected gram positive bacteria and this
property is enhanced if further purified by HPLC. The materials
being treated to provide inhibition are preferably foods, although
other materials can be treated.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX

Sequence 321 BP; 114 A; 51 C; 63 G; 93 T; 0 other;

Query Match 41.6%; Score 319.4; DB 14; Length 321;
Best Local Similarity 99.7%; Pred. No. 2.9e-55;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

198 CCGGAGCAATTAATTAAGGCTCTGATTAATTCGAGTTGTTAGTACAAATGATTCGT 257
DB 1 CCGGAGCAATTAATTAAGGCTCTGATTAATTCGAGTTGTTAGTACAAATGATTCGT 60

258 TCGAAGCACTCAAAATTAATTAAGAGGCACTCAAAATGAGTACAAAGATTTTAA 317
DB 61 TCGAAGCACTCAAAATTAATTAAGAGGCACTCAAAATGAGTACAAAGATTTTAA 120

318 CTTCGATTTGGTATCTGTTTCGAAGAAAGATTCAAGTCACACGATTAACAAGTAT 377
DB 121 CTTCGATTTGGTATCTGTTTCGAAGAAAGATTCAAGTCACACGATTAACAAGTAT 180

378 TTTCGATTTGGTATCTGTTTCGAAGAAAGATTCAAGTCACACGATTAACAAGTAT 437
DB 181 TTTCGATTTGGTATCTGTTTCGAAGAAAGATTCAAGTCACACGATTAACAAGTAT 240

438 AACTTGTCAATTTGATTTACGTTAAGCAAAATTAACCAATCAAGAGATGATTTTGTGA 497
DB 241 AACTTGTCAATTTGATTTACGTTAAGCAAAATTAACCAATCAAGAGATGATTTTGTGA 300

498 GTTCAGACATGATTAATCTATCC 518
DB 301 GTTCAGACATGATTAATCTATCC 321

RESULT 10
AAQ34782
ID AAQ34782 standard; DNA; 321 BP.
XX AAQ34782;
AC
XX
XX 25-MAR-2003 (updated)
DT 12-MAY-1993 (first entry)

XX Bacteriocin Ll-2 precursor gene.
DE
XX Food treatment; ss.
KM
XX
XX Lactococcus lactis sub-species lactis NRRL B-18809.
OS

Key Location/Qualifiers
CDS 101..274
/*tag= a
FT sig_peptide 101..169
FT /*tag= b
FT mat_peptide 170..271
FT /*tag= c

US5173297-A.
22-DEC-1992.
01-JUL-1991; 91US-0721774.
01-JUL-1991; 91US-0721774.
(UNIL) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
Vedamuthu ER, Henderson JT, Marugg JD, Vanmassenaar PD;
WPI; 1993-017533/02.
P-PSDB; AAR33850.

Inhibition of Gram-positive bacteria - using bacteriocin derived
from *Lactococcus lactis* sub-species *lactis* NRRL B-18809

Disclosure; Page 13; 14pp; English.

The sequence is that of the bacteriocin Ll-2 precursor gene, the
precursor can be used in a method for the inhibition of Gram-positive
bacteria. Ll-2 is especially useful for treatment of food, although
other non-food materials may also be treated.
CC (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX

Sequence 321 BP; 114 A; 51 C; 63 G; 93 T; 0 other;

Query Match 41.6%; Score 319.4; DB 14; Length 321;
Best Local Similarity 99.7%; Pred. No. 2.9e-55;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

198 CCGGAGCAATTAATTAAGGCTCTGATTAATTCGAGTTGTTAGTACAAATGATTCGT 257
DB 1 CCGGAGCAATTAATTAAGGCTCTGATTAATTCGAGTTGTTAGTACAAATGATTCGT 60

258 TCGAAGCACTCAAAATTAATTAAGAGGCACTCAAAATGAGTACAAAGATTTTAA 317
DB 61 TCGAAGCACTCAAAATTAATTAAGAGGCACTCAAAATGAGTACAAAGATTTTAA 120

318 CTTCGATTTGGTATCTGTTTCGAAGAAAGATTCAAGTCACACGATTAACAAGTAT 377
DB 121 CTTCGATTTGGTATCTGTTTCGAAGAAAGATTCAAGTCACACGATTAACAAGTAT 180

378 TTTCGATTTGGTATCTGTTTCGAAGAAAGATTCAAGTCACACGATTAACAAGTAT 437
DB 181 TTTCGATTTGGTATCTGTTTCGAAGAAAGATTCAAGTCACACGATTAACAAGTAT 240

438 AACTTGTCAATTTGATTTACGTTAAGCAAAATTAACCAATCAAGAGATGATTTTGTGA 497
DB 241 AACTTGTCAATTTGATTTACGTTAAGCAAAATTAACCAATCAAGAGATGATTTTGTGA 300

498 GTTCAGACATGATTAATCTATCC 518
DB 301 GTTCAGACATGATTAATCTATCC 321

RESULT 11
AAQ46818
ID AAQ46818 standard; DNA; 321 BP.
XX
XX AAQ46818;
AC
XX 25-MAR-2003 (updated)
DT 21-JAN-1994 (first entry)
XX
XX Lactococcal bacteriocin coding sequence.
DE
XX Bacteriocin; inhibition; polypeptide; Lactococcus lactis; ss.
KW
XX Lactococcus lactis (subspecies lactis).
OS
XX
XX Key Location/Qualifiers
FH CDS 101..271
FT /*tag= a
FT /product= Bacteriocin.
FT /note= "The bacteriocin comprises the 34 terminal
FT amino acids of the polypeptide precursor."
XX
XX US5231165-A.
PN
XX 27-JUL-1993.
PD
XX 14-MAY-1992; 92US-0882715.
PF
XX 01-JUL-1991; 91US-0721774.
PR 14-MAY-1992; 92US-0882715.
XX
XX (UNCL) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
PA
XX Henderson JT, Marugs JD, Vanwassenaar PD, Vedamuthu ER;
PI
XX MPI; 1993-249768/31.
DR P-PSDB; AAR39312.
XX
XX Isolated and purified polypeptide from Lactococcus lactis sub
PT species lactis - has inhibitory activity against gram-positive
PT bacteria for e.g. food etc.
XX
XX Claim 1; Column 17-18; 13pp; English.
PS
XX The isolated bacteriocin obtained from the polypeptide precursor has
CC an inhibitory activity against selected Gram positive bacteria.
CC The amount of bacteriocin required to provide inhibition is 15-100
CC arbitrary units per gram of material. The materials being treated
CC with the bacteriocin to provide inhibition are especially foodstuffs.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC
XX Sequence 321 BP; 114 A; 51 C; 63 G; 93 T; 0 other;
SQ
Query Match 41.6%; Score 319.4; DB 14; Length 321;
Best Local Similarity 99.7%; Pred. No. 2.9e-55;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 198 CGCGAGCAATATTAACGGCTCGATTAATTCGTAAGTTGTTAGATCAATGATTTGCT 257
Db 1 CGCGAGCAATATTAACGGCTCGATTAATTCGTAAGTTGTTAGATCAATGATTTGCT 60
Oy 258 TCGAAGGAATCAAAATTAATTAAGAGGCACTCAAAATGAGTACAAAAGATTTTA 317
Db 61 TCGAAGGAATCAAAATTAATTAAGAGGCACTCAAAATGAGTACAAAAGATTTTA 120
Oy 318 CTTCGATTTGATCTGTTTGAAGAAGATTCAAGTGCATACACGCGATTACAAGTAT 377
Db 121 CTTCGATTTGATCTGTTTGAAGAAGATTCAAGTGCATACACGCGATTACAAGTAT 180
Oy 378 TTCCGTATGTACACCCGGTTGTAAACAGAGCTCTGATGGGTTGTAAATGAAAAACGC 437
Db 181 TTCCGTATGTACACCCGGTTGTAAACAGAGCTCTGATGGGTTGTAAATGAAAAACGC 240

Oy 438 AACTTCATTTGATGATTCACGTACGAATTAACCAATCAAGATAGTATTTGTTA 497
Db 241 AACTTCATTTGATGATTCACGTACGAATTAACCAATCAAGATAGTATTTGTTA 300
Oy 498 GTTCAGACATGATACATATCC 518
Db 301 GTTCAGACATGATACATATCC 321

RESULT 12
AA169229
ID AA169229 standard; DNA; 212 BP.
XX
XX AA169229;
AC
XX 07-FEB-2002 (first entry)
DT
XX Plasmid pMTL910E DNA fragment.
DE
XX Plasmid pMTL910E; integration; ss.
KW
XX Unidentified.
OS
XX
XX Key Location/Qualifiers
FH -35_signal 71..77
FT /*tag= a
FT -10_signal 94..100
FT /*tag= b
XX
XX WO20017319-A2.
PN
XX 18-OCT-2001.
PD
XX 09-APR-2001; 2001WO-GB01612.
PF
XX 07-APR-2000; 2000GB-0008682.
PR
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA
XX Minton NP, Purdy DA, Elmore MJ, O'Keefe KMT;
PI
XX MPI; 2002-010908/01.
DR
XX
XX New plasmid, designated pCD6, useful for transformation of Clostridium
PT difficile and for expressing gene in Clostridium difficile
PT
XX
XX Example 3; Fig 12; 60pp; English.
PS
XX This invention describes a novel plasmid (1), designated pCD6 for
CC transformation of Clostridium difficile. The invention describes a method
CC 1) for (a) expressing a heterologous gene sequence in C. difficile which
CC comprises providing a plasmid containing the heterologous gene sequence,
CC introducing the plasmid into C. difficile, and optionally, where the
CC plasmid also contains a gene coding for a selectable marker, selecting
CC for C. difficile that express the selectable marker; (2) for making a
CC plasmid for expression of a heterologous coding sequence in C. difficile
CC comprises: (3) for identifying a C. difficile virulence factor which
CC comprises culturing C. difficile in the absence of, and in the presence
CC of a regulating factor that promotes expression of C. difficile virulence
CC factors and identifying a putative virulence factor whose expression is
CC reduced in the absence of the regulating factor compared with the
CC expression in the presence of the regulating factor; (4) for identifying
CC a vector that integrates into a gram positive bacterial genome which
CC comprises: (a) transforming a gram positive bacteria with a plasmid,
CC where the plasmid comprises an inducible promoter and replication of the
CC plasmid is dependent upon presence of an inducer of the promoter, where
CC the plasmid includes a sequence coding for a selectable marker, and the
CC transformation takes place in the presence of the inducer, removing the
CC inducer, and selecting for bacteria expressing the selectable marker; or
CC (b) transforming a gram positive bacteria with a plasmid, where the
CC plasmid comprises an suppressible promoter and replication of the plasmid
CC is dependent upon absence of a suppressor of the promoter, where the

CC plasmid includes a sequence coding for a selectable marker, and the
 CC transformation takes place in the absence of the suppressor, adding the
 CC suppressor, and selecting for bacteria expressing the selectable marker;
 CC (5) for identifying a C. difficile methylase gene which comprises
 CC identifying the sequence of a bacterial methylase gene, comparing the
 CC sequence of the gene with the genome of a strain of C. difficile,
 CC identifying a region of the genome that contains at least 30% homology
 CC with the bacterial methylase gene, and expressing that region; (1) is
 CC useful for expressing a gene in C. difficile by making (1) containing
 CC that gene and transforming C. difficile with (1); (1) is useful for
 CC transforming C. difficile. Method (4) is useful for identification of a
 CC vector that integrates into Clostridia. This sequence a DNA fragment from
 CC the plasmid pMTL910E which is used to illustrate the method of the
 CC invention.

SO Sequence 212 BP; 83 A; 31 C; 33 G; 65 T; 0 other;

Query Match 26.4%; Score 202.4; DB 24; Length 212;
 Best Local Similarity 97.2%; Pred. No. 8.9e-32;
 Matches 206; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 92 GATGACCTGATCTTAACTAATGACATGACATGAAACATTAACATCTAAACGCTCT 151
 DB 1 GATGACCTGATCTTAACTAATGACATGACATGAAACATTAACATCTAAACGCTCT 60
 QY 152 TAATCTATCTGAGAAAGTATGTAATATTAATTTGTCGATAACGAGCATATAA 211
 DB 61 TAATCTATCTGAGAAAGTATGTAATATTAATTTGTCGATAACGAGCATATAA 120
 QY 212 ACGGCTCTGATTAATTTGTAAGTTGTTAGATACAAATGTTGTCGAAGCACTACA 271
 DB 121 ACGGCTCTGATTAATTTGTAAGTTGTTAGATACAAATGTTGTCGAAGCACTACA 180
 QY 272 AATATAATTATAAGAGGACCTCAAAATGAGT 303
 DB 181 AATATAATTATAAGAGGACCTCAAAATGAGT 212

RESULT 13
 AA050289

ID AA050289 standard; DNA; 546 BP.

XX AA050289;
 AC 25-MAR-2003 (updated)
 DT 03-MAY-1994 (first entry)
 XX
 XX Lactose operon promoter.
 DE
 XX Nisin; antibiotic; food; spoilage; pathogens; industrial processes;
 KM lantibiotic; Lactococcus lactis; ss.
 XX
 XX Escherichia coli.
 OS
 XX WO9320213-A1.
 PN
 XX 14-OCT-1993.
 PD
 XX 01-APR-1993; 93WO-GB00676.
 PF
 XX 02-APR-1992; 92GB-0007267.
 PR
 XX (AGRI-) AGRIC & FOOD RES COUNCIL.
 PA
 XX Dodd HM, Gasson MJ;
 FI
 XX WPI; 1993-336920/42.
 DR
 XX
 XX
 PT New organisms able to express genes for nisin maturation, but not
 PT native pre-nisin - useful for producing variant nisin(s) for
 PT antimicrobial use in food industry
 XX
 XX Disclosure; Figure 5; 62pp; English.

XX Nisin is a highly modified peptide antibiotic produced by certain
 CC bacteria e.g. strains of Lactococcus lactis. It has efficient
 CC antimicrobial activity against a wide range of Gram positive
 CC bacteria which are food pathogens or spoilage organisms. The genes
 CC involved in nisin maturation can be used to transform other microbes
 CC which are unable to secrete the natural nisin nisin. These organisms
 CC can however modify nisin and translocate it out of the cell. This
 CC provides an effective means of producing variant ninsins which are
 CC useful as antimicrobial agents in conditions of high pH where
 CC natural nisin is ring-opened and loses its antimicrobial properties.
 CC Variant nisa genes were placed under the regulation of the lactose
 CC operon promoter and used to transform cells.
 CC (Updated on 25-MAR-2003 to correct PN field.)

SO Sequence 546 BP; 229 A; 62 C; 78 G; 177 T; 0 other;

Query Match 11.1%; Score 84.8; DB 14; Length 546;
 Best Local Similarity 97.7%; Pred. No. 4.1e-08;
 Matches 86; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 213 CGGCTCTGATTAATTTGAAAGTTGTTAGATACAAATGATTCGTAAGCACTACA 272
 DB 459 CGGATCCGATTAATTTGAAAGTTGTTAGATACAAATGATTCGTAAGCACTACA 518
 QY 273 AATAAATTATAAGAGGACCTCAAAATG 300
 DB 519 AATAAATTATAAGAGGACCTCAAAATG 546

RESULT 14

AAx87829 standard; DNA; 144 BP.

XX AAx87829;
 AC 09-NOV-1999 (first entry)
 DT
 XX
 DE DNA encoding subtilin-nisin chimera.
 XX
 XX Nisin; subtilin; lantibiotic; chimera; mutant; bacteriocide;
 KM preservative; ss.
 XX
 XX Chimeric - Lactococcus lactis.
 OS Chimeric - Bacillus subtilis.
 XX
 XX Key Location/Qualifiers
 FH CDS 1..126
 FT /*tag= a
 FT 1..21
 FT /*tag= b
 FT /note= "subtilin signal peptide"
 FT 22..123
 FT /*tag= c
 FT /note= "Sub (1-11)-Nis (12-32) "
 XX
 XX WO9711713-A1.
 PN
 XX 03-APR-1997.
 PD
 XX 30-SEP-1996; 96WO-US15160.
 PF
 XX 28-SEP-1995; 95US-0535494.
 PR
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA
 XX Hansen JN;
 PI
 XX
 XX
 DR WPI: 1997-225847/20.
 DR P-PSDB; AAY31659.
 XX
 XX Lantibiotic mutants and chimera(s) - having enhanced stability and
 PT activity compared to nisin

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 11:23:51 ; Search time 67 Seconds
(without alignments)
5052.849 Million cell updates/sec

Title: US-10-082-618-8

Perfect score: 767

Sequence: 1 agtcgacgaatcattcaataa.....tcgacgaatcattcgcgtcgt 767

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	675	88.0	7454	4	US-08-836-687B-29
2	440	57.4	448	3	US-08-773-731A-1
3	410.4	53.5	1446	2	US-08-560-007B-4
4	358.4	46.7	360	1	US-08-129-151A-1
5	358.4	46.7	360	2	US-08-715-579-1
6	319.4	41.6	321	1	US-07-721-774A-1
7	319.4	41.6	321	1	US-07-882-715-1
8	319.4	41.6	321	1	US-07-882-079-1
9	300	39.1	300	4	US-08-836-687B-19
10	296.8	38.7	300	4	US-08-836-687B-22
11	293.6	38.3	300	4	US-08-836-687B-21
12	293.6	38.3	300	4	US-08-836-687B-23
13	293.6	38.3	300	4	US-08-836-687B-25
14	290.4	37.9	300	4	US-08-836-687B-27
15	84.8	11.1	546	3	US-08-773-731A-3
16	81.8	10.7	144	2	US-08-535-494-8
17	81.8	10.7	144	2	US-09-097-635-8
18	53.6	7.0	545	1	US-07-981-525-6
19	53.6	7.0	545	1	US-08-220-033-6
20	53.6	7.0	545	1	US-08-465-491-6
21	53.6	7.0	545	2	US-08-986-617-6
22	52.8	6.9	100	3	US-08-773-731A-7
23	52	6.8	545	1	US-07-981-525-8
24	52	6.8	545	1	US-08-220-033-8
25	52	6.8	545	2	US-08-465-491-8
26	52	6.8	545	2	US-08-986-617-8
27	50.4	6.6	545	1	US-07-981-525-1

28	50.4	6.6	545	1	US-08-220-033-1	Sequence 1, Appl1
29	50.4	6.6	545	2	US-08-465-491-1	Sequence 1, Appl1
30	50.4	6.6	545	2	US-08-986-617-1	Sequence 1, Appl1
31	49.6	6.5	100	3	US-08-773-731A-21	Sequence 21, Appl1
32	48.8	6.4	545	1	US-07-981-525-3	Sequence 3, Appl1
33	48.8	6.4	545	1	US-08-220-033-3	Sequence 3, Appl1
34	48.8	6.4	545	2	US-08-465-491-3	Sequence 3, Appl1
35	48.8	6.4	545	2	US-08-986-617-3	Sequence 3, Appl1
36	44.6	5.8	7218	1	US-08-232-463-14	Sequence 14, Appl1
37	44.6	5.8	19124	2	US-08-487-826B-13	Sequence 13, Appl1
38	44.4	5.8	127	2	US-08-535-494-4	Sequence 4, Appl1
39	44.4	5.8	127	3	US-09-097-635-4	Sequence 4, Appl1
40	43.6	5.7	8654	1	US-08-920-812-6	Sequence 6, Appl1
41	43.6	5.7	8654	1	US-08-920-812-6	Sequence 6, Appl1
42	43.6	5.7	8654	1	US-08-921-177-6	Sequence 6, Appl1
43	43.6	5.7	8654	1	US-08-362-577C-6	Sequence 6, Appl1
44	43.6	5.7	8654	2	US-08-920-828-6	Sequence 6, Appl1
45	42.8	5.6	1826	3	US-09-286-691-11	Sequence 11, Appl1

ALIGNMENTS

RESULT 1									
US-08-836-687B-29									
Sequence 29, Application US/08836687B									
Patent No. 6448034									
GENERAL INFORMATION:									
APPLICANT: Gasson, Michael John									
APPLICANT: Dodd, Helen Mair									
TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN									
FILE REFERENCE: 20747/70									
CURRENT APPLICATION NUMBER: US/08/836,687B									
CURRENT FILING DATE: 1995-11-20									
NUMBER OF SEQ ID NOS: 51									
SOFTWARE: PatentIn Ver. 2.1									
SEQ ID NO 29									
LENGTH: 7454									
TYPE: DNA									
ORGANISM: Lactococcus sp.									
US-08-836-687B-29									
Query Match									
Best Local Similarity 100.0%; Score 675; DB 4; Length 7454;									
Matches 675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	93	AATGACCTAGTCTTATTAATACTGACCAATAGAAACATTAACAATCTTAAGAGTCTT	152						
DB	1	AATGACCTAGTCTTATTAATACTGACCAATAGAAACATTAACAATCTTAAGAGTCTT	60						
QY	153	AATTCATCTTGGAAGAAGTATGGTAAATATATTTGTGATTAACGGAGCATTAATAA	212						
DB	61	AATTCATCTTGGAAGAAGTATGGTAAATATATTTGTGATTAACGGAGCATTAATAA	120						
QY	213	CGGCTCGATTAAATTCGAAAGTTGTAGATACCAATGATTCGTCGAAGAACTACAA	272						
DB	121	CGGCTCGATTAAATTCGAAAGTTGTAGATACCAATGATTCGTCGAAGAACTACAA	180						
QY	273	AATAAATTAATAGAGGACCTCAAAATAGATGACAAAAGATTTAAGTTGGATTC	332						
DB	181	AATAAATTAATAGAGGACCTCAAAATAGATGACAAAAGATTTAAGTTGGATTC	240						
QY	333	TGTTTCGAGAAAGATTCAGGTGCATCACCGCATTTCAAGTATTTGGTATGTACACC	392						
DB	241	TGTTTCGAGAAAGATTCAGGTGCATCACCGCATTTCAAGTATTTGGTATGTACACC	300						
QY	393	CGGTTGTAACAGAGAGCTCGATGGGTGTAACATGAAACAGCAACTTGTCAATGTAG	452						
DB	301	CGGTTGTAACAGAGAGCTCGATGGGTGTAACATGAAACAGCAACTTGTCAATGTAG	360						
QY	453	TATTCAGTGAAGCAAAATTAACCAATCAAGAGATATTTGTTAGTTAGTACAACTGATA	512						
DB	361	TATTCAGTGAAGCAAAATTAACCAATCAAGAGATATTTGTTAGTTAGTACAACTGATA	420						

QY 513 CTATCTATTTTATTAAGTTATTTAGGTTGCTAAATAGCTATATAAAATTAAGAGAGA 572
| | | | |
Db 421 CTATCTATTTTATTAAGTTATTTAGGTTGCTAAATAGCTATATAAAATTAAGAGAGA 480
| | | | |
QY 573 AAAACATGATATAAAAGTTCAATTTAAAGCTCAACCGTTTATTAAGTAAGAAATCAATTTTA 632
| | | | |
Db 481 AAAACATGATATAAAAGTTCAATTTAAAGCTCAACCGTTTATTAAGTAAGAAATCAATTTTA 540
| | | | |
QY 633 TCTCCAAAGATTAACGGAGTTTACTGAATATATCTCAAGTCAATGAGCTGAAGTTAA 692
| | | | |
Db 541 TCTCCAAAGATTAACGGAGTTTACTGAATATATCTCAAGTCAATGAGCTGAAGTTAA 600
| | | | |
QY 693 AATAAAGTTTTTTGGAAAGTTACTAGCTAAGCTTAATCTTAAGCTTAATGATGTTAGCAG 752
| | | | |
Db 601 AATAAAGTTTTTTGGAAAGTTACTAGCTAAGCTTAATCTTAAGCTTAATGATGTTAGCAG 660
| | | | |
QY 753 AATAAATATGCTGCT 767
| | | | |
Db 661 AATAAATATGCTGCT 675
| | | | |

RESULT 2
US-08-773-731A-1
Sequence 1, Application US/08773731A
Patent No. 6100056

GENERAL INFORMATION:

APPLICANT: Gasson, Michael J.
APPLICANT: Dodd, Helen M.
TITLE OF INVENTION: NISINS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP
STREET: 2101 L Street N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/773,731A
FILING DATE: 24-DEC-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/313,123
FILING DATE: 18-NOV-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00676
FILING DATE: 01-APR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9207267.7
FILING DATE: 02-APR-1992

ATTORNEY/AGENT INFORMATION:
NAME: Brady, Jr., James W.
REGISTRATION NUMBER: 32,115
REFERENCE/DOCKET NUMBER: E8280.016/P016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-785-9700
TELEFAX: 202-887-0689

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: CDNA
FEATURE:

NAME/KEY: CDS
LOCATION: join(88..258, 369..446)

US-08-773-731A-1

Query Match 57.4%; Score 440; DB 3; Length 448;
Best Local Similarity 98.2%; Pred. No. 9,8e-89;
Matches 440; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 211 AACGGCTGATTAATTAATCTGAAGTTGTTAGATACAAATGATTTGTTGAAGAACTAC 270
| | | | |
Db 1 AACGGTNGANTTAATTTCTGAAGTTGTTAGATACAAATGATTTGTTGAAGAACTAC 60
| | | | |
QY 271 AAAATTAATTAAGAGGCACTCAAAAATGATACAAAAGATTTTAACCTGGATTGGTA 330
| | | | |
Db 61 AAAATTAATTAAGAGGCACTCAAAAATGATACAAAAGATTTTAACCTGGATTGGTA 120
| | | | |
QY 331 TCTGTTGGAAGAAAGATTCAGTGCTACACACGCAATTAAGATTTGCTATGATCA 390
| | | | |
Db 121 TCTGTTGGAAGAAAGATTCAGTGCTACACACGCAATTAAGATTTGCTATGATCA 180
| | | | |
QY 391 CCCGGTTGTAACAGAGAGCTCTGATGGTTGTATACATGAACAGCAACTGTCTATTGT 450
| | | | |
Db 181 CCCGGTTGTAACAGAGAGCTCTGATGGTTGTATACATGAACAGCAACTGTCTATTGT 240
| | | | |
QY 451 AGTATTCAGTAAGCAAAATTAACCAATCAAGAGATGATTTGTTAGTTGACATGCA 510
| | | | |
Db 241 AGTATTCAGTAAGCAAAATTAACCAATCAAGAGATGATTTGTTAGTTGACATGCA 300
| | | | |
QY 511 TACTATCTATTTTATTAAGTTATTTAGGTTGCTAAATAGCTTAATAAAATTAAGAGAG 570
| | | | |
Db 301 TACTATCTATTTTATTAAGTTATTTAGGTTGCTAAATAGCTTAATAAAATTAAGAGAG 360
| | | | |
QY 571 GAAAAAACATGATATAAAAGTTCAATTTAAAGCTCAACCGTTTATTAAGTAAGAAATCAATTT 630
| | | | |
Db 361 GAAAAAACATGATATAAAAGTTCAATTTAAAGCTCAACCGTTTATTAAGTAAGAAATCAATTT 420
| | | | |
QY 631 TATCTCAACGATTAACGAGGTTTAC 658
| | | | |
Db 421 TATCTCAACGATTAACGAGGTTTAC 448
| | | | |

RESULT 3

US-08-560-007B-4
Sequence 4, Application US/08560007B
Patent No. 5914248

GENERAL INFORMATION:

APPLICANT: KUIPERS, OSCAR PAUL
APPLICANT: DE VOS, WILHEM MEINDERT
TITLE OF INVENTION: METHOD FOR CONTROLLING THE GENE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE WEBB LAW FIRM
STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE
CITY: PITTSBURGH
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 15219-1818

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,007B
FILING DATE: 17 NOV 1995
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1446
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN

US-08-560-007B-4

Query Match	53.5%	Score 410.4	DB 2	Length 1446
Best Local Similarity	99.8%	Pred. No. 4e-82		
Matches 411	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 1	AGTTGACGAATATTTAATAATTTTATTAATATCTTGATTTTCTAGTTCCTGAATAAATATA	60		
Db 1035	AGTTGACGAATATTTAATAATTTTATTAATATCTTGATTTTCTAGTTCCTGAATAAATATA	1094		
QY 61	GAGATAGCTTATTAGAGCTTAGACATPACTGGAATGACCTAGCTTAACTAATACAGAC	120		
Db 1095	GAGATAGCTTATTAGAGCTTAGACATPACTGGAATGACCTAGCTTAACTAATACAGAC	1154		
QY 121	AATGAAACATTAAACAAATCTAAAAAGCTTTAATTCATCTGTGGAAGAATGGTAT	180		
Db 1155	AATGAAACATTAAACAAATCTAAAAAGCTTTAATTCATCTGTGGAAGAATGGTAT	1214		
QY 181	AATATTTATTTGCGATTAACGCGAGCATTAATAACGGCTCTGAATTAATTTCTAAGTTGTT	240		
Db 1215	AATATTTATTTGCGATTAACGCGAGCATTAATAACGGCTCTGAATTAATTTCTAAGTTGTT	1274		
QY 241	AGATACATATGATTTGCTTCGAAAGGAACATCAAAAATTAATTTAAGGAGGCACTCAAAATG	300		
Db 1275	AGATACATATGATTTGCTTCGAAAGGACATCAAAAATTAATTTAAGGAGGCACTCAAAATG	1334		
QY 301	AGTACAAAAGATTTTAACTTGATTTGGTATCTGTTTGAAGAAAGATTCAGGTGCATCA	360		
Db 1335	AGTACAAAAGATTTTAACTTGATTTGGTATCTGTTTGAAGAAAGATTCAGGTGCATCA	1394		
QY 361	CCAGAGCATTTACAAGATTTGCGCTATGTATACACCGGTTCTTAATAACAGAGGCTC	412		
Db 1395	CCAGAGCATTTACAAGATTTGCGCTATGTATACACCGGTTCTTAATAACAGAGGCTC	1446		

RESULT 4
US-08-129-151A-1
Sequence 1, Application US/08129151A
Patent No. 5594103
GENERAL INFORMATION:
APPLICANT: DE VOS, Willem M.
APPLICANT: SIEZEN, Roelant J.
APPLICANT: KUIPERS, Oscar P.
TITLE OF INVENTION: LANTIBIOTICS SIMILAR TO NISIN A, LACTIC
TITLE OF INVENTION: ACID BACTERIA WHICH PRODUCE SUCH LANTIBIOTICS, METHOD FOR
TITLE OF INVENTION: CONSTRUCTING SUCH LACTIC ACID BACTERIA AND METHOD FOR
TITLE OF INVENTION: PRESERVING FOODSTUFFS WITH ETC.
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: c/o YOUNG & THOMPSON
STREET: 745 South 23rd Street, Second Floor
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/129,151A
FILING DATE: 07-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NL 9100634
FILING DATE: 11-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00068
FILING DATE: 09-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: BO 37078
TELECOMMUNICATION INFORMATION:

```

? TELEPHONE: 703/521-2297
? TELEFAX: 703/685-0573
? TELEX: 248425 EMBON
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 360 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 121..294
?
US-08-129-151A-1

Query Match          46.7%   Score 358.4; DB 1; Length 360;
Best local Similarity 99.7%   Pred. No. 9,5e-71;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0

```

```

Query Match          46.7%; Score 358.4; DB 1; Length 360;
Best local similarity 99.7%; Pred. No. 9.5e-71;
Matches 35; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      178 AATAATATTATTGTGCGATACGCGAGCATATATAACGCGCTGATTAAATTCGAA GTT 237
DB      1 AATAATATTATTGTGCGATACCGCGAGCATATAAAGCGCTGATTAAATTCGAA GTT 60

QY      238 GTTAGTATAAATGATTTTCGTTGGAAGAACTACAAATTAATTAAGAGGCACTGCAA 297
DB      61 GTTAGATACAAATGATTTTCGTTGGAAGAACTACAAATTAATTAAGAGGCACTGCAA 120

QY      298 ATGAGTACAAAAGATTTTAACTTTGATTTGATCTGTTTCGAGAAAGATTCAAGTGCA 357
DB      121 ATGAGTACAAAAGATTTTAACTTTGATTTGATCTGTTTCGAGAAAGATTCAAGTGCA 180

QY      358 TCACCAAGCATTAACAAGATTTTCGATATGTACACCGGTTGTAAAAAGAGACTTGATG 417
DB      181 TCACCAAGCATTAACAAGATTTTCGATATGTACACCGGTTGTAAAAAGAGACTTGATG 240

QY      418 GGTGTGAACATGAAAACAGCACTTGTCATTTAGTATTACGTAAAGCAATTAACCAAT 477
DB      241 GGTGTGAACATGAAAACAGCACTTGTCATTTAGTATTACGTAAAGCAATTAACCAAT 300

QY      478 CAAGAGATGTATTTTGTAGTTGACAGACATGATACCTATTTTAATAGTAAATTA 537
DB      301 CAAGAGATGTATTTTGTAGTTGACAGACATGATACCTATTTTAATAGTAAATTA 360

RESULT 5
US-08-715-579-1
; Sequence 1, Application US/08715579
; Patent No. 5928946
; GENERAL INFORMATION:
; APPLICANT: DE VOS, Willem M.
; APPLICANT: SIEZEN, Roelant J.
; APPLICANT: KUIPERS, Oscar P.
; TITLE OF INVENTION: LANTIBIOTICS SIMILAR TO NISIN A, LACTIC
; TITLE OF INVENTION: ACID BACTERIA WHICH PRODUCE SUCH LANTIBIOTICS, METHOD FOR
; TITLE OF INVENTION: CONSTRUCTING SUCH LACTIC ACID BACTERIA AND METHOD FOR
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: C/O YOUNG & THOMPSON
; STREET: 745 South 23rd Street, Second Floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/715,579
; FILING DATE: 19-SEP-1996
; CLASSIFICATION: 435

```

Query Match	Best Local Similarity	46.7%	Score 358.4	DB 2	Length 360;
Matches 359;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	178	AATATATTTATTTGTCGATMACGCGAGCATATTAACGGCTCTGATTAATTCGAAAGTT	237		
Db	1	AATATATTTATTTGTCGATMACGCGAGCATATTAACGGCTCTGATTAATTCGAAAGTT	60		
QY	238	GTTTNGATACAAATGATTTTCGTCGAAAGAACTACAAAATTAATTTATTAAGAGGCACTCAA	297		
Db	61	GTTTNGATACAAATGATTTTCGTCGAAAGAACTACAAAATTAATTTATTAAGAGGCACTCAA	120		
QY	298	ATGAGTACAAAGATTTTAACTTGATTTGGTATCTGTTTGAAGAAAGATTCAAGTGA	357		
Db	121	ATGAGTACAAAGATTTTAACTTGATTTGGTATCTGTTTGAAGAAAGATTCAAGTGA	180		
QY	358	TCACACGCGATTACAGATATTCGCTATGACACCGGGTTGTAACAGAGACTCTGATG	417		
Db	181	TCACACGCGATTACAGATATTCGCTATGACACCGGGTTGTAACAGAGACTCTGATG	240		
QY	418	GGTGTGAACATGAAACAGCAACTTGTCATTTGTATGATTCAGTAAGCAATTAACCAAT	477		
Db	241	GGTGTGAACATGAAACAGCAACTTGTCATTTGTATGATTCAGTAAGCAATTAACCAAT	300		
QY	478	CAAGGATAGTATTTTGTAGTTCAGACATGATACATCTATTTTATTAAGTTATTTA	537		
Db	301	CAAGGATAGTATTTTGTAGTTCAGACATGATACATCTATTTTATTAAGTTATTTA	360		

RESULT 6

US-07-721-774A-1

; Sequence 1, Application US/07721774A

; Patent No. 5173297

GENERAL INFORMATION:

APPLICANT: Vedamuthu, Ebenezer R; Henderson, James T;

APPLICANT: Marang, John D; van Wassenar, Plietier D

TITLE OF INVENTION: No. 5173297a1 Bacteriocin From *Lactococcus lactis*

TITLE OF INVENTION: Subspecies *lactis*

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ian C. McLeod

STREET: 2190 Commons Parkway

CITY: Okemos

Query Match	41.6%	Score 319.4	DB 1	Length 321
Best Local Similarity	99.7%	Pred. No. 3,8e-62		
Matches 320	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 198	CGGAGCATATATAAACGGCTCTGATTAAATTCGAAATTGTTAGATACATGATTTGGT	257		
Db 1	CGGAGCATATATAAACGGCTCTGATTAAATTCGAAATTGTTAGATACATGATTTTCGT	60		
QY 258	TGCAAGGAAGCTACAAAATTAATATATATAGAGGAGCATCAAAATAGATACAAAAGATTTTAA	317		
Db 61	TCGAGAGAACTACAAAATTAATATATATAGAGGAGCATCAAAATAGATACAAAAGATTTTAA	120		
QY 318	CTTGAATTTGGTATCTGTTTCGAAGAAAGATTGAGGTGATCAACAGCATTTCAAGTAT	377		
Db 121	CTTGAATTTGGTATCTGTTTCGAAGAAAGATTGAGGTGATCAACAGCATTTCAAGTAT	180		
QY 378	TTGGCTATGTACACCCGGTTGTAAACAGAGAGCTCGATGGCTTTGAAACATGAAACAGC	437		
Db 181	TTGGCTATGTACACCCGGTTGTAAACAGAGAGCTCGATGGCTTTGAAACATGAAACAGC	240		
QY 438	AACCTGTCAATTTAGATATTCACGTAAAGCAAAATACCAAAATCAAAAGATAGTATTTGTAA	497		
Db 241	AACCTGTCAATTTAGATATTCACGTAAAGCAAAATACCAAAATCAAAAGATAGTATTTGTAA	300		
QY 498	GTTCAAGACATGATACTATCC 518			
Db 301	GTTCAAGACATGATACTATCC 321			

RESULT 7
US-07-882-715-1
; Sequence 1, Application US/07882715
; Patent No. 5231165

GENERAL INFORMATION:
APPLICANT: Ebenezer R. Vedamuthu, James T.
APPLICANT: Henderson, John D. Marug, Pieter D.
APPLICANT: van Wassenaer
TITLE OF INVENTION: No. 5231165e1 Bacteriocin From
TITLE OF INVENTION: Lactococcus lactis subspecies
TITLE OF INVENTION: lactis
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
COMPUTER: Acer
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/882,715
FILING DATE: 19920514
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/721,774
FILING DATE: July 1, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MT 4.1-142
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 5231165e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: Other DNA
HYPOTHETICAL: No
ANTI-SENSE: No
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis subspecies lactis
STRAIN: NRRL-B-18809
INDIVIDUAL ISOLATE: LLA-2.0
CELL TYPE: Unicellular organism
FEATURE:
NAME/KEY: bacteriocin encoding DNA
LOCATION: 101 to 271
IDENTIFICATION METHOD: sequencing
US-07-882-715-1

Query Match 41.6%; Score 319.4; DB 1; Length 321;
Best Local Similarity 99.7%; Pred. No. 3.8e-62;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

198 CGCAGCATATAAAGCGCTCGATTAATTCGAAAGTTGTTAGATACATGATTTGCT 257
1 CGCGAGCATATAAAGCGCTCGATTAATTCGAAAGTTGTTAGATACATGATTTGCT 60
258 TCGAAGAACTACAAATTAATTAAGAGGCACTCAAAATGAGTACAAAAGATTTTAA 317
61 TCGAAGAACTACAAATTAATTAAGAGGCACTCAAAATGAGTACAAAAGATTTTAA 120
318 CTTGATTTGGATCTGTTTCCAGAAAGATTCAGGTGATACACGCGATTTCAAGTAT 377
121 CTTGATTTGGATCTGTTTCCAGAAAGATTCAGGTGATACACGCGATTTCAAGTAT 180
378 TTCGATATACACCGGTTGTAACAGAGCTGATGGTGTGATGATGATGATGATGATGAT 437

181 TTCGATATACACCGGTTGTAACAGAGCTGATGGTGTGATGATGATGATGATGATGAT 240
438 AACTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 497
241 AACTGTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
498 GTTCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 518
301 GTTCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 321

RESULT 8
US-07-882-079-1
Sequence 1, Application US/07882079
Patent No. 5232849
GENERAL INFORMATION:
APPLICANT: Ebenezer R. Vedamuthu, James T.
APPLICANT: Henderson, John D. Marug, Pieter D.
APPLICANT: van Wassenaer
TITLE OF INVENTION: No. 5232849e1 Bacteriocin From
TITLE OF INVENTION: Lactococcus lactis subspecies
TITLE OF INVENTION: lactis
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
COMPUTER: Acer
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/882,079
FILING DATE: 19920514
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/721,774
FILING DATE: July 1, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MT 4.1-144
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
TELEX: No. 5232849e
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: Other DNA
HYPOTHETICAL: No
ANTI-SENSE: No
ORIGINAL SOURCE:
ORGANISM: Lactococcus lactis subspecies lactis
STRAIN: NRRL-B-18809
INDIVIDUAL ISOLATE: LLA-2.0
CELL TYPE: Unicellular organism
FEATURE:
NAME/KEY: bacteriocin encoding DNA
LOCATION: 101 to 271
IDENTIFICATION METHOD: sequencing
US-07-882-079-1

Query Match 41.6%; Score 319.4; DB 1; Length 321;
Best Local Similarity 99.7%; Pred. No. 3.8e-62;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 198 CCGGACATATTAACGGCTCTGATTAAATTCGAAGTTGTTAGATACATGATTTCGT 257
DB 1 CCGGACATATTAACGGCTCTGATTAAATTCGAAGTTGTTAGATACATGATTTCGT 60
QY 258 TCGAAGAACTACAAATAATTAATTAAGAGGACCTCAAAATGATGCAAAAGATTTTAA 317
DB 61 TCGAAGAACTACAAATAATTAATTAAGAGGACCTCAAAATGATGCAAAAGATTTTAA 120
QY 318 CTGGATTGTTGATCTGTTTCCGAAGAAAGATTCAAGGTCACCAAGATTCAAGTAT 377
DB 121 CTGGATTGTTGATCTGTTTCCGAAGAAAGATTCAAGGTCACCAAGATTCAAGTAT 180
QY 378 TTGCTATGTACACCGGTTGTTAAACAGAGCTCTGATGGTTGTATGATGAAAACAGC 437
DB 181 TTGCTATGTACACCGGTTGTTAAACAGAGCTCTGATGGTTGTATGATGAAAACAGC 240
QY 438 AACTTGTCAATTTAGATTAATTCAGCTAAGCAATAATCAAAATGATGATTTTGTGA 497
DB 241 AACTTGTCAATTTAGATTAATTCAGCTAAGCAATAATCAAAATGATGATTTTGTGA 300
QY 498 GTTCAGACATGATCTATCC 518
DB 301 GTTCAGACATGATCTATCC 321

```

RESULT 9

```

US-08-836-687B-19
; Sequence 19, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Lactococcus sp.
US-08-836-687B-19

```

```

Query Match 39.1%; Score 300; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 7.1e-58;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 207 AATAAAGCGCTCTGATTAAATTCGAAGTTGTTAGATACATGATTTCGTTGAAGAA 266
DB 1 AATAAAGCGCTCTGATTAAATTCGAAGTTGTTAGATACATGATTTCGTTGAAGAA 60
QY 267 CTACAAATAATTAATTAAGAGGACCTCAAAATGATGCAAAAGATTTTAATTGATTT 326
DB 61 CTACAAATAATTAATTAAGAGGACCTCAAAATGATGCAAAAGATTTTAATTGATTT 120
QY 327 GGTATCTGTTTGAAGAAAGATTCAAGTGCATCACAGCATTACAAAGATTTCGCTATG 386
DB 121 GGTATCTGTTTGAAGAAAGATTCAAGTGCATCACAGCATTACAAAGATTTCGCTATG 180
QY 387 TACACCCGGTTGTTAAACAGAGCTCTGATGGTTGTATGATGAAAACAGCAATTTGTA 446
DB 181 TACACCCGGTTGTTAAACAGAGCTCTGATGGTTGTATGATGAAAACAGCAATTTGTA 240
QY 447 TTGATGATTCAGTGAAGCAATAATCAAAATGATGATGATTTTGTAGTTGACAGACA 506
DB 241 TTGATGATTCAGTGAAGCAATAATCAAAATGATGATGATTTTGTAGTTGACAGACA 300

```

```

RESULT 10
US-08-836-687B-22
; Sequence 22, Application US/08836687B

```

```

; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Lactococcus sp.
US-08-836-687B-22

```

```

Query Match 38.7%; Score 296.8; DB 4; Length 300;
Best Local Similarity 99.3%; Pred. No. 3.6e-57;
Matches 298; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 207 AATAAAGCGCTCTGATTAAATTCGAAGTTGTTAGATACATGATTTCGTTGAAGAA 266
DB 1 AATAAAGCGCTCTGATTAAATTCGAAGTTGTTAGATACATGATTTCGTTGAAGAA 60
QY 267 CTACAAATAATTAATTAAGAGGACCTCAAAATGATGCAAAAGATTTTAATTGATTT 326
DB 61 CTACAAATAATTAATTAAGAGGACCTCAAAATGATGCAAAAGATTTTAATTGATTT 120
QY 327 GGTATCTGTTTGAAGAAAGATTCAAGTGCATCACAGCATTACAAAGATTTCGCTATG 386
DB 121 GGTATCTGTTTGAAGAAAGATTCAAGTGCATCACAGCATTACAAAGATTTCGCTATG 180
QY 387 TACACCCGGTTGTTAAACAGAGCTCTGATGGTTGTATGATGAAAACAGCAATTTGTA 446
DB 181 TACACCCGGTTGTTAAACAGAGCTCTGATGGTTGTATGATGAAAACAGCAATTTGTA 240
QY 447 TTGATGATTCAGTGAAGCAATAATCAAAATGATGATGATTTTGTAGTTGACAGACA 506
DB 241 TTGATGATTCAGTGAAGCAATAATCAAAATGATGATGATTTTGTAGTTGACAGACA 300

```

RESULT 11

```

US-08-836-687B-21
; Sequence 21, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Lactococcus sp.
US-08-836-687B-21

```

```

Query Match 38.3%; Score 293.6; DB 4; Length 300;
Best Local Similarity 98.7%; Pred. No. 1.8e-56;
Matches 296; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 207 AATAAAGCGCTCTGATTAAATTCGAAGTTGTTAGATACATGATTTCGTTGAAGAA 266
DB 1 AATAAAGCGCTCTGATTAAATTCGAAGTTGTTAGATACATGATTTCGTTGAAGAA 60
QY 267 CTACAAATAATTAATTAAGAGGACCTCAAAATGATGCAAAAGATTTTAATTGATTT 326
DB 61 CTACAAATAATTAATTAAGAGGACCTCAAAATGATGCAAAAGATTTTAATTGATTT 120
QY 327 GGTATCTGTTTGAAGAAAGATTCAAGTGCATCACAGCATTACAAAGATTTCGCTATG 386

```


Db 121 GGTATCTGTTGGAAGAAAGATTCAGTGCATCACACGCATTAAGATTTCCCTATG 180
Qy 387 TACACCCGGTGTAAACAGAGCTCTGATGGTTGTACATGAAGAAACGCAACTTGTCA 446
Db 181 TACACCCGGTGTAAACAGAGCTCTGATGGTTGTACATGAAGAAACGCAACTTGTCA 240
Qy 447 TTGTAGTATTCAGTAAGCAATTAACCAATCAAGATAGTATTTTGTAGTTCAGACA 506
Db 241 TTGTAGTATTCAGTAAGCAATTAACCAATCAAGATAGTATTTTGTAGTTCAGACA 300

RESULT 12

US-08-836-687B-23
; Sequence 23, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Lactococcus sp.
US-08-836-687B-23

Query Match 38.3%; Score 293.6; DB 4; Length 300;

Best Local Similarity 98.7%; Pred. No. 1.8e-56;
Matches 296; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 207 AATAACGGCTCTGATTAATTTCTGAAGTTGTAGATCAATGATTTGTTGGAAGAA 266
Db 1 AATAACGGATCGATTAATTTCTGAAGTTGTAGATCAATGATTTGTTGGAAGAA 60
Qy 267 CTACAAATTAATTAATTAAGAGGCACTCAAAATGATGACAAAGATTTTAACCTTGATTT 326
Db 61 CTACAAATTAATTAATTAAGAGGCACTCAAAATGATGACAAAGATTTTAACCTTGATTT 120
Qy 327 GGTATCTGTTGGAAGAAAGATTCAGTGCATCACACGCATTAAGATTTCCCTATG 386
Db 121 GGTATCTGTTGGAAGAAAGATTCAGTGCATCACACGCATTAAGATTTCCCTATG 180
Qy 387 TACACCCGGTGTAAACAGAGCTCTGATGGTTGTACATGAAGAAACGCAACTTGTCA 446
Db 181 TACACCCGGTGTAAACAGAGCTCTGATGGTTGTACATGAAGAAACGCAACTTGTCA 240
Qy 447 TTGTAGTATTCAGTAAGCAATTAACCAATCAAGATAGTATTTTGTAGTTCAGACA 506
Db 241 TTGTAGTATTCAGTAAGCAATTAACCAATCAAGATAGTATTTTGTAGTTCAGACA 300

RESULT 13

US-08-836-687B-25
; Sequence 25, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Lactococcus sp.

US-08-836-687B-25

Query Match 38.3%; Score 293.6; DB 4; Length 300;

Best Local Similarity 98.7%; Pred. No. 1.8e-56;
Matches 296; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 207 AATAACGGCTCTGATTAATTTCTGAAGTTGTAGATCAATGATTTGTTGGAAGAA 266
Db 1 AATAACGGATCGATTAATTTCTGAAGTTGTAGATCAATGATTTGTTGGAAGAA 60
Qy 267 CTACAAATTAATTAATTAAGAGGCACTCAAAATGATGACAAAGATTTTAACCTTGATTT 326
Db 61 CTACAAATTAATTAATTAAGAGGCACTCAAAATGATGACAAAGATTTTAACCTTGATTT 120
Qy 327 GGTATCTGTTGGAAGAAAGATTCAGTGCATCACACGCATTAAGATTTCCCTATG 386
Db 121 GGTATCTGTTGGAAGAAAGATTCAGTGCATCACACGCATTAAGATTTCCCTATG 180
Qy 387 TACACCCGGTGTAAACAGAGCTCTGATGGTTGTACATGAAGAAACGCAACTTGTCA 446
Db 181 TACACCCGGTGTAAACAGAGCTCTGATGGTTGTACATGAAGAAACGCAACTTGTCA 240
Qy 447 TTGTAGTATTCAGTAAGCAATTAACCAATCAAGATAGTATTTTGTAGTTCAGACA 506
Db 241 TTGTAGTATTCAGTAAGCAATTAACCAATCAAGATAGTATTTTGTAGTTCAGACA 300

RESULT 14

US-08-836-687B-27
; Sequence 27, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Lactococcus sp.
US-08-836-687B-27

Query Match 37.9%; Score 290.4; DB 4; Length 300;

Best Local Similarity 98.0%; Pred. No. 9.3e-56;
Matches 294; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 207 AATAACGGCTCTGATTAATTTCTGAAGTTGTAGATCAATGATTTGTTGGAAGAA 266
Db 1 AATAACGGATCGATTAATTTCTGAAGTTGTAGATCAATGATTTGTTGGAAGAA 60
Qy 267 CTACAAATTAATTAATTAAGAGGCACTCAAAATGATGACAAAGATTTTAACCTTGATTT 326
Db 61 CTACAAATTAATTAATTAAGAGGCACTCAAAATGATGACAAAGATTTTAACCTTGATTT 120
Qy 327 GGTATCTGTTGGAAGAAAGATTCAGTGCATCACACGCATTAAGATTTCCCTATG 386
Db 121 GGTATCTGTTGGAAGAAAGATTCAGTGCATCACACGCATTAAGATTTCCCTATG 180
Qy 387 TACACCCGGTGTAAACAGAGCTCTGATGGTTGTACATGAAGAAACGCAACTTGTCA 446
Db 181 TACACCCGGTGTAAACAGAGCTCTGATGGTTGTACATGAAGAAACGCAACTTGTCA 240
Qy 447 TTGTAGTATTCAGTAAGCAATTAACCAATCAAGATAGTATTTTGTAGTTCAGACA 506
Db 241 TTGTAGTATTCAGTAAGCAATTAACCAATCAAGATAGTATTTTGTAGTTCAGACA 300

RESULT 15

US-08-773-731A-3

```

: Sequence 3 Application US/08/773731A
: Patent No. 6100056
:
: GENERAL INFORMATION:
:
: APPLICANT: Gasson, Michael J.
: APPLICANT: Dodd, Helen M.
: TITLE OF INVENTION: NISINS
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP
: STREET: 2101 L Street N.W.
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20037
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/773,731A
: FILING DATE: 24-DEC-1996
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/313,123
: FILING DATE: 18-NOV-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/GB93/00676
: FILING DATE: 01-APR-1993
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9207267.7
: FILING DATE: 02-APR-1992
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brady, Jr., James W.
: REGISTRATION NUMBER: 32,115
: REFERENCE/DOCKET NUMBER: E8280.016/P016
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-785-9700
: TELEFAX: 202-887-0689
:
: INFORMATION FOR SEQ ID NO: 3:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 546 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: DNA (genomic)
:
: IS-08-773-731A-3

```

	Query Match	Similarity	Score	DB	Length
Best Local	86%	97.7%	84.8	DB 3	546
Matches	86	Conservative	0	Mismatches	2
				Indels	0
				Gaps	0

Search completed: January 12, 2004, 12:06:14
Job time : 69 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	767	100.0	767	13	US-10-082-618-8	Sequence 8, Appli
2	534	7.0	11905	13	US-10-311-455-1721	Sequence 1721, Ap
3	50	6.5	3673778	13	US-10-312-641-1	Sequence 1, Appli
4	48.8	6.4	6361	13	US-10-311-455-1113	Sequence 1113, Ap
5	48.4	6.3	13125	13	US-10-311-455-1200	Sequence 1200, Ap
6	48.4	6.3	13125	13	US-10-240-445-110	Sequence 110, App
7	48.2	6.3	4985	15	US-10-056-440-10	Sequence 10, Appl
8	48.2	6.3	4985	15	US-10-094-240-10	Sequence 10, Appl
9	48	6.3	17594	13	US-10-311-455-1999	Sequence 1999, Ap
10	47.2	6.2	7351	13	US-10-311-455-2	Sequence 2, Appli
11	47	6.1	5452	13	US-10-311-455-1122	Sequence 1122, Ap
12	47	6.1	3673778	13	US-10-312-641-2	Sequence 2, Appl
13	46.8	6.1	5303	13	US-10-311-455-84	Sequence 84, App
14	46.8	6.1	11294	13	US-10-311-455-959	Sequence 959, App
15	46.6	6.1	7403	13	US-10-311-455-2189	Sequence 2189, Ap

C 16	46.6	6.1	13123	13	US-10-240-452-64	Sequence 64, Appl
C 17	46.6	6.1	3673778	13	US-10-312-851-22	Sequence 2, Appl
C 18	46.4	6.0	16127	13	US-10-311-455-718	Sequence 718, Appl
C 19	46.4	6.0	40862	13	US-10-311-455-2045	Sequence 2045, Appl
C 20	46	6.0	5689	13	US-10-240-453-100	Sequence 100, Appl
C 21	46	6.0	5689	15	US-10-239-676-90	Sequence 90, Appl
C 22	45.8	6.0	5743	13	US-10-311-455-2041	Sequence 2041, Appl
C 23	45.8	6.0	5987	13	US-10-311-455-1536	Sequence 1536, Appl
C 24	45.8	6.0	13814	13	US-10-311-455-1166	Sequence 1166, Appl
C 25	45.6	5.9	5511	13	US-10-311-455-1843	Sequence 1843, Appl
C 26	45.4	5.9	6120	13	US-10-240-453-313	Sequence 313, Appl
C 27	45.4	5.9	6904	13	US-10-311-455-185	Sequence 185, Appl
C 28	45.2	5.9	5324	13	US-10-311-455-1763	Sequence 1763, Appl
C 29	45.2	5.9	16287	13	US-10-311-455-646	Sequence 646, Appl
C 30	44.8	5.8	17594	13	US-10-311-455-1999	Sequence 1999, Appl
C 31	44.6	5.8	755	13	US-10-027-632-128109	Sequence 128109, Appl
C 32	44.6	5.8	755	14	US-10-027-632-128109	Sequence 128109, Appl
C 33	44.6	5.8	10716	13	US-10-311-455-1392	Sequence 1392, Appl
C 34	44.4	5.8	19021	13	US-10-311-455-1943	Sequence 1943, Appl
C 35	44.4	5.8	10006	13	US-10-311-455-9	Sequence 9, Appl11
C 36	44.4	5.8	13574	13	US-10-311-455-1290	Sequence 1290, Appl
C 37	44.4	5.8	18624	13	US-10-311-455-1676	Sequence 1676, Appl
C 38	44.2	5.8	6029	13	US-10-311-455-1966	Sequence 1966, Appl
C 39	44.2	5.8	8801	13	US-10-311-455-1714	Sequence 1714, Appl
C 40	44.2	5.8	8801	13	US-10-240-453-160	Sequence 160, Appl
C 41	44.2	5.8	8801	15	US-10-239-676-144	Sequence 144, Appl
C 42	44	5.7	8467	13	US-10-311-455-82	Sequence 82, Appl
C 43	44	5.7	12405	13	US-10-240-453-43	Sequence 43, Appl
C 44	44	5.7	12405	15	US-10-239-676-35	Sequence 35, Appl
C 45	43.8	5.7	2000	10	US-09-938-842A-5203	Sequence 5203, Appl

ALIGNMENTS

```

RESULT 1
US-10-082-618-8
; Sequence 8, Application US/10082618
; Publication No. US20030175207A1
;
GENERAL INFORMATION:
;
APPLICANT: OLSTEIN, ALAN D.
APPLICANT: PEITAG, JOELEN
TITLE OF INVENTION: BACTERIOICIN-METAL COMPLEXES IN THE DETECTION OF
; TITLE OF INVENTION: PATHOGENS AND OTHER BIOLOGICAL ANALYTES
; FILE REFERENCE: 7005-0003
;
CURRENT APPLICATION NUMBER: US/10/082,618
;
CURRENT FILING DATE: 2002-02-22
;
NUMBER OF SEQ ID NOS: 8
;
SOFTWARE: PatentIn Ver. 2.1
;
SEQ ID NO 8
;
LENGTH: 767
;
TYPE: DNA
;
ORGANISM: Streptococcus lactis
;
US-10-082-618-8

```

Query Match	100.0%;	Score 767;	DB 13;	Length 767;
Best Local Similarity	100.0%;	Pred. No. 3.6e-148;		
Matches 767; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

[illegible]

Db 181 AATATATATGCGATACCGGACATATAAAGCGCTGATTAATTCGAAATTGTT 240
Qy 241 AATATCAAGATTTCTGCGAAGAACTCAAAATTAATTAAGAGGACCTCAAAAG 300
Db 241 AATATCAAGATTTCTGCGAAGAACTCAAAATTAATTAAGAGGACCTCAAAAG 300
Qy 301 AGTACAAAAGATTTTAACTTGATTTGATCTGTTCCAGAAAGATTCAGTGCA 360
Db 301 AGTACAAAAGATTTTAACTTGATTTGATCTGTTCCAGAAAGATTCAGTGCA 360
Qy 361 CCAACGATTAACAATTTGCTGATTAACCCGGTTGTAACAGAGCTCTGATGGT 420
Db 361 CCAACGATTAACAATTTGCTGATTAACCCGGTTGTAACAGAGCTCTGATGGT 420
Qy 421 TGTAAATGAAAACAGCACTGCTGATTAATTCAGTAAGCAAAATTAACAAATCA 480
Db 421 TGTAAATGAAAACAGCACTGCTGATTAATTCAGTAAGCAAAATTAACAAATCA 480
Qy 481 AGGATAGATTTTGTAGTTGACAGATGATCTATCTATTTTAAATTTAGAGG 540
Db 481 AGGATAGATTTTGTAGTTGACAGATGATCTATCTATTTTAAATTTAGAGG 540
Qy 541 TTGCTAAATAGCTTAAATTAAGAGAGAAATTAAGATTAAGATTTTAAAG 600
Db 541 TTGCTAAATAGCTTAAATTAAGAGAGAAATTAAGATTAAGATTTTAAAG 600
Qy 601 CTCACCGTTTTAGTAAAGAAATTAATTAATCTCAACGATTAACGAGTTTACTG 660
Db 601 CTCACCGTTTTAGTAAAGAAATTAATTAATCTCAACGATTAACGAGTTTACTG 660
Qy 661 AATATCTCAAGCTGATGAGCTGATTAAGATTAAGATTTTGAAGAGTTACTAC 720
Db 661 AATATCTCAAGCTGATGAGCTGATTAAGATTAAGATTTTGAAGAGTTACTAC 720
Qy 721 TAGCTAATCTTAACTCTATGATGATTAACAGAAATTAATGCTGCT 767
Db 721 TAGCTAATCTTAACTCTATGATGATTAACAGAAATTAATGCTGCT 767

RESULT 2
US-10-311-455-1721
; Sequence 1721, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIERENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1721
; LENGTH: 11805
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1721

Query Match 7.0%; Score 53.4; DB 13; Length 11805;
Best Local Similarity 47.2%; Pred. No. 0.61;
Matches 162; Conservative 0; Mismatches 181; Indels 0; Gaps 0;
12 ATTATATATTTATTAATCTGATTTTCTAGTTCTGATTAATTAAGATGAGTTT 71

Db 1032 ATTATAGATTTTATTAAGATTTGATTAATTTGATTAATTAATTAATGAGGTTT 1091
Qy 72 ATTAGCTTTAGACACTACTGATTAAGACCTAGCTTAATTAAGTACATTAAGAACT 131
Db 1092 ATTATTAATAGTTTATTAAGGAGAAATTAATTAATTTGATTAATTTGTTAATG 1151
Qy 1132 TAACAAATCTAAACAGCTTAAATCTTCTGAGAAATTAATTTGATTAATTTGTT 191
Db 1132 TTAGAAATGATTTTAAATGATTTTGGTTTAAAGTATGATTTTAAATTAATTAAGT 1211
Qy 192 CGATTAACCGAGCATTAATTAACCGCTGATTAATTTGATTAATTTGATTAATGA 251
Db 1212 AATATTAAGAGGAGATTAATTAAGATTTGATTTTAAATTTTAAAGGATTAATTA 1271
Qy 252 TTTGCTGGAAGAACTCAAAATTAATTAAGAGGACCTCAAAATGAGTAAAGAA 311
Db 1272 TGTAGTTATATGTTGTTTAAATTAATTAATTAAGAGGAAATTTGAGGAGGAGGAA 1331
Qy 312 TTTTAACTTGGATTTGATCTGTTTCGAGAAAGATTCAGGT 354
Db 1332 TTTGAGTTTGGTTTGGGATTTTAAATTTTATTTTGGGT 1374

RESULT 3
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenome AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 6.5%; Score 50; DB 13; Length 3673778;
Best Local Similarity 48.9%; Pred. No. 29;
Matches 134; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 483 GATGATATTTGTTGTTGACATGATGATCTATCTTTTAAATTAATTTAGGTT 542
Db 2686866 GTTAAGAGTTATTAATTAATTAATTTATTTATGATTAATGCAAGTTGTTATTTT 2686755
Qy 543 GCTAAATAGCTTAAATTAAGAGAGAAATTAAGATTAATTAATTAATTAAGCT 602
Db 2686756 TTGTAGAAATTAATTAAGAGAGAAATTAATTAATTAATTAATTAAGTAAAGTA 2686815
Qy 603 CAACGTTTTTATGTAAGAAATTAATTAATTTATCTCAAAAGATTAACGAGTTTACTGAA 662
Db 2686816 AATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2686875
Qy 663 TATACTCAAGCATTAAGAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 722
Db 2686876 TTGTAGAGATTAAG 2686935
Qy 723 GCTAATCTTAACCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 756
Db 2686936 GAAATTAATTAATTAATTAATTTTATTTATTAATTAATTAATTAATTAATTAATTA 2686969

RESULT 4
US-10-311-455-1113
; Sequence 1113, Application US/10311455

```
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1113
; LENGTH: 6361
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1113
Query Match      6.4%; Score 48.8; DB 13; Length 6361;
Best Local Similarity 49.1%; Pred. No. 4.2; Mismatches 162; Indels 4; Gaps 1;
Matches 160; Conservative 0;

Qy      1 AGTTGACCAATATTTAATTAATTTAATTAATCTTGATTTTCTAGTCCGAAATATATA 60
Db      1315 AGTGAATGATATATATATGATGATTTTATAGGAATGATTTTATGTTTATGTTTATAC 1374

Qy      61 GAGATAGTTTATGAGCTTACGATACCTTGAAAGACCTGCTTAACTATATCTGAC 120
Db      1375 GTGATTTATAGAAAAGTAAATTTAGTTATGAAAATGTAAATAATATATAAT 1434

Qy      121 AATGAAACATTAACAATCTTAAACAGCTTAATCTATCTGAGAAAGTATGTAAT 180
Db      1435 ATTGTATATAGAAAAGTAAATATATATATATATATATATATATATATATATAT 1494

Qy      181 AATATTTATGTCGATTAACGCGAGCATATTAACCGCTCGATTAATCTGAAGTTGTT 240
Db      1495 TAATTTTGTAATAATATATATATATATATATATATATATATATATATATATAT 1550

Qy      241 AGATACAAATGATTTGCTTGAAGCACTACAAATTAATTAAGAGGCACTCAAAATG 300
Db      1551 ATTATTTAGTTATTTGATGAGGAGGTAATATATATATTTTATGTAAGAAATATAATG 1610

Qy      301 AGTACAAAGATTTTAATCTGATTT 326
Db      1611 GCTTTAAATTTTAAATTTTATGATTT 1636

RESULT 5
US-10-311-455-1200/C
; Sequence 1200, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
```

```
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1200
; LENGTH: 13125
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 11070, 11234
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1200
Query Match      6.3%; Score 48.4; DB 13; Length 13125;
Best Local Similarity 46.9%; Pred. No. 6.8; Mismatches 171; Indels 0; Gaps 0;
Matches 151; Conservative 0;

Qy      421 TGTACATGAAAACAGCACTTGTCAATTTAGTATTCAGTACCAATTAACCAATCA 480
Db      2139 TATATAAACAATCTACCAATTAACCTTAATTAATCAATCAATTAATTAATTAAT 2080

Qy      481 AGATAGATTTTGTAGTTGACATGATACATATCTATTTTATTAAGTTATTTAGG 540
Db      2079 ACATTACTTAAACCAAAAACAAAATTAATTAATTTAATTAATTAATTAATTAAT 2020

Qy      541 TTGCTAATAGCTTATTAATAAATAAGAGAGAAAACATGATTAATTAATTAATTAAG 600
Db      2019 TTATTTAATTTATTAACCACTTAACAAATTAATAAATAAATAAATAAATAAATAA 1960

Qy      601 CTCACCGTTTATTAAGTAAGAAATCAATTTTATCTCAACAGATTAACGAGTTTACTG 660
Db      1959 CACGAAAATTCATCAACATTAACCAATTAATAACAATCAACCAATCAATTTTATTA 1900

Qy      661 AATTTACTCAAGTATGAGCTGTAAGTAATAAATTAAGTTTATTTGGAACATTTACTAC 720
Db      1899 CACTTAATCTAATATATATATATATATATATTAATTAATTAATTTCTAATAATTCAAACCTCT 1840

Qy      721 TAGTAATCCTAACCTCATGA 742
Db      1839 CAACCTATCAAACTCTCATTA 1818

RESULT 6
US-10-240-485-110/C
; Sequence 110, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240.485
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 110
; LENGTH: 13125
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
```

```

; NAME/KEY: unsure
; LOCATION: (11070, 11234)
US-10-240-485-110

```

Query Match	6.3%;	Score 48.4;	DB 13;	Length 13125;
Best Local Similarity	46.9%;	Pred. No. 6.8;		
Matches 151;	Conservative 0;	Mismatches 171;	Indels 0;	Gaps 0;

Qy	421	TGTAACTGAAAAACGCAACTGTGATGTAGTATCCAGTACAAATTAACCAATGCA	480
Db	2139	TATAAAACAAATCTACAAATTAACCTTAATTAATAATCTCCAAATTAATACAAATTTA	2086
Qy	481	AGGATGATATTTGGTATGTTGAGACATGATATCTATCCATTTTATATAGTTATTTAGG	540
Db	2079	ACATTACTTAAACCAAAAAACAAAAAATTAATTAATTTAAATTAATCATATATCTCA	2022
Qy	541	TTGCTAAATAGCTTATATAAAAAATAAGAGGAGAAAAACATGATATAAAAGTTCAITTTAAAG	600
Db	2019	TTATTAATCTTATTAACCACTTAACAAATTAATTAATAAATAAATAAATAAATCTCAAACTAATA	1966
Qy	601	CTCAACCGCTTTTGTATTAAGAAATATACATTTTATCTCCAAAGCATTAACGAGAGTTTACTG	660
Db	1959	CACGAAAAATTCATATCAACATATACCAATTAACCAACTCAACAAAAATTCGATTTTCTTA	1900
Qy	661	AATATATCTCAAGTCATTTGAGACTGTGAATTAATAAAGTTTTTTTGGACAGTTACTAC	720
Db	1899	CACCTAACTAATATATATATATATATACATTTAAATAAATAATTTCTAATAAATTCAAACTCT	1840
Qy	721	TAGCTAATCTTAAACTCTATGCA	742
Db	1839	CAACTATCAAACTCTCTATATA	1818

```

RESULT 7
US-10-056-405-10/C
; Sequence 10, Application US/10056405
; Publication No. US20030166013A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: N7841
; CURRENT APPLICATION NUMBER: US/10/056,405
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4985
; TYPE: DNA
; ORGANISM: Anopheles gambiae
; US-10-056-405-10

```

Query Match	6.3%	Score 48.2	DB 13	Length 4985
Best Local Similarity	47.0%	Pred. No. 5.1		
Matches 149	Conservative 0	Mismatches 166	Indels 0	Gaps 0
QY	1	AGTTGACGAATATTTAATAATTTATTATTAATCTTGATTTTCTGATTCCTGAATATATATA	60	
Db	2680	AATTAATTAATAAATAAATAAATAAATAAATAATATTTATTTGTTAATAATAAATAAATAATAT	2621	
QY	61	GAGATAGGTTATTGAGCTTAGACATACCTGGAATGACCTAGTCTTAATCACTATACGAC	120	
Db	2620	AATATATTAATAATAAT	2561	
QY	121	AATGAAACATTAACAATCTAAACACTCTTAATTCTATCTTGAAAGATTTGGTAAT	180	
Db	2560	AAT	2501	
QY	181	AATATTATTGTCGATTAACGCGACATATAAACGGCTCTGATTAATTCTGAAGTTGTT	240	
Db	2500	AAT	2441	

QY 241 AGATACCATGATGTTTCGTCGAAGCACTACAAATTAATTTAAGGAGGCCTCAAAATG 300
Db 2440 AATTAT 2381
QY 301 AGTACAAAAGATTTTAA 317
Db 2380 AATTATATATATATATATTA 2364

```

RESULT 8
US-10-094-240-10/c
: Sequence 10. Application US/10094240
: Publication No. US20030082637A1
: GENERAL INFORMATION:
: APPLICANT: ZWIEBEL, LAURENCE J
: TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
: FILE REFERENCE: N8289
: CURRENT APPLICATION NUMBER: US/10/094,240
: CURRENT FILING DATE: 2001-03-08
: PRIOR APPLICATION NUMBER: 10/056,405
: PRIOR FILING DATE: 2002-01-24
: PRIOR APPLICATION NUMBER: 60/264,649
: PRIOR FILING DATE: 2001-01-26
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 10
: LENGTH: 4985
: TYPE: DNA
: ORGANISM: Anopheles gambiae
US-10-094-240-10

```

Query Match	Best Local Similarity	6.3% ; Score 48.2; DB 15; Length 4985;
Matches 149;	Conservative 0;	Mismatches 168; Indels 0; Gaps 0;
Qy	1	AGTTGACGAATATTTAATAATTATTATTAATATCTGATTTTCTAGTCTCGAATTAATATA 60
Db	2680	AATTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2621
Qy	61	GAGATAGGTTTATTGAGCTTTGACATACCTTGAATGACCTAGCTTTAACTAATACGAC 120
Db	2620	AATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2561
Qy	121	AATGAACAACATTAACAATCTPAACACGCTTAATCTCATCTTGAGAAAGATTTGTAAT 180
Db	2360	AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2501
Qy	181	AATATTATTGTCGATPACGGGAGCATATAACCGGCTTCGATTAATTCGAAAGCTTTGTT 240
Db	2500	AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2441
Qy	241	AGATACAAATGATTTTCGTTCCGAGGAACATCAAAAATAATTATTAAGGAGGCGCTCAAAATG 300
Db	2440	AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2381
Qy	301	AGTCAAAAGATTTTAA 317
Db	2380	AATTAATTAATTAATTAATTA 2364

```

RESULT 9
US-10-311-455-1999
; Sequence 1999, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation Status of Cytosine Methylation
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455

```

;; CURRENT FILING DATE: 2002-12-16
;; PRIOR APPLICATION NUMBER: PCT/EP01/07537
;; PRIOR FILING DATE: 2001-07-02
;; PRIOR APPLICATION NUMBER: DE 10032529.7
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: DE 10043826.1
;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 2424
;; SEQ ID NO 1999
;; LENGTH: 17594
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1999

Query Match 6.3%; Score 48; DB 13; Length 17594;
Best Local Similarity 47.9%; Pred. No. 9.2;
Matches 138; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 475 AATCAAGAGTATGTTTGTAGTTCAGACATGATCTATCTATTTTAAAGTAT 534
DB 15916 AATTTAAAT 15975
QY 535 TTAGGTTGCTAATAGCTTATATAAATTAAGAGAGAGAGAGAGAGAGAGAGAG 594
DB 15976 TTATAGTATTTTAAATATATATATATATATATATATATATATATATATAT 16035
QY 595 TTAAGCTCAACCGTTTGTAGTGAAGATATCAATTTTATCTCAACGATTAACGAGTT 654
DB 16036 TTTTATAGTATGATGTTGTGTGAAGATATTTTATTTTATTTTATTTTAT 16095
QY 655 TTACGATATATCTCAAGCATGAGCTGATAGTAAATATATATATATATATAT 714
DB 16096 AAT 16155
QY 715 TACTACTAGCTAATCTTAAACTCTATGATGTTATGACAGAAATATATATG 762
DB 16156 TAAAAATATAGAAATATTTGAAATTTAGAGTTTAAAAATATATATATTTG 16203

RESULT 10
US-10-311-455-2/c
;; Sequence 2, Application US/10311455
;; Publication No. US20030143606A1
;; GENERAL INFORMATION:
;; APPLICANT: OLEK, Alexander
;; APPLICANT: PIPENBROCK, Christian
;; APPLICANT: BERLIN, Kurt
;; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
;; FILE REFERENCE: 5013.1014
;; CURRENT APPLICATION NUMBER: US/10/311,455
;; PRIOR FILING DATE: 2002-12-16
;; PRIOR APPLICATION NUMBER: PCT/EP01/07537
;; PRIOR FILING DATE: 2001-07-02
;; PRIOR APPLICATION NUMBER: DE 10032529.7
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: DE 10043826.1
;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 2424
;; SEQ ID NO 2
;; LENGTH: 7351
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2

Query Match 6.2%; Score 47.2; DB 13; Length 7351;
Best Local Similarity 48.8%; Pred. No. 9.5;
Matches 160; Conservative 0; Mismatches 163; Indels 5; Gaps 1;

QY 9 AATATTAATAG 68
DB 1286 AATATTAAT 1227
QY 69 TTTATGAGCTTATAGACATATGATGACCTAGTCTTATATATATATATATATAT 123
DB 1226 TTTACCTACCCACATATATATATATATATATATATATATATATATATATAT 1167
QY 124 AGAATATTAAT 183
DB 1166 ATATATTAAT 1107
QY 184 ATATATGATTAAT 243
DB 1106 ATATATGATTAAT 1047
QY 244 TACATATGATTTGTTGAGAGACTACAAATATATATATATATATATATATAT 303
DB 1046 CATATAATATCAATTTCAAAAAAAAAAAAAAAAAAAAAAAAAATATATATCA 987
QY 304 ACAAAGATTTTACTGATTTGGTAT 331
DB 986 AAAAAATTTTAT 959

RESULT 11
US-10-311-455-1122
;; Sequence 1122, Application US/10311455
;; Publication No. US20030143606A1
;; GENERAL INFORMATION:
;; APPLICANT: OLEK, Alexander
;; APPLICANT: PIPENBROCK, Christian
;; APPLICANT: BERLIN, Kurt
;; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
;; FILE REFERENCE: 5013.1014
;; CURRENT APPLICATION NUMBER: US/10/311,455
;; PRIOR FILING DATE: 2002-12-16
;; PRIOR APPLICATION NUMBER: PCT/EP01/07537
;; PRIOR FILING DATE: 2001-07-02
;; PRIOR APPLICATION NUMBER: DE 10032529.7
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: DE 10043826.1
;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 2424
;; SEQ ID NO 1122
;; LENGTH: 5452
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1122

Query Match 6.1%; Score 47; DB 13; Length 5452;
Best Local Similarity 44.3%; Pred. No. 9.3;
Matches 286; Conservative 0; Mismatches 350; Indels 9; Gaps 2;

QY 4 TGACGAATATTAAT 63
DB 3747 TAAAAAT 3806
QY 64 ATAGTTATATGATCTTATAGACATATGATGATGATGATGATGATGATGATGAT 123
DB 3807 ATTTGTTAT 3866
QY 124 AGAATATTAAT 183
DB 3867 TTTATTTTAT 3926
QY 184 ATATATGATTAAT 243
DB 3927 ATTTTGGAGCTGATGAGAGGATATATATATATATATATATATATATATAT 3986

Oy	244	TACATGATTTTCGTCGAAGCACTACAAAATTAATTTAAAGAGCCCTCAAAATAGT	303
Db	3887	AATATGGAAATTTTGTGTTATTTAAAAATTAATATTTAGTTGGCGCTGTGTAGGT	4046
Oy	304	ACAAAGATTTTAACTTGATTTGGTATCTGTTTCGAAGAAAGATTCT-----AGGTG	355
Db	4047	ATTATATATTTTAGTTATTTTAGAGGTTGAGATAGAAATCGTTTGAATTCGGAGGCG	4106
Oy	356	CATCACCCGCAATTACAAAGTATTCGCTATGTACCCGGTTGTAAAAACAGAGCTCTGA	415
Db	4107	GAGGTTGAGTGAGATTAAAGATTATATTAATTGTATATTTAGTTGGGTGTAAGAGAGAAAT	4166
Oy	416	TGGGTTGPAACATGAAACAGCAACTGTCAATTGAGATTACGTAAGCAAAATACCAA	475
Db	4167	TTGTGTTTAAAAATBAATBAATBAATBAATBAATBAATBAATBAATBAATBAATBA	4226
Oy	476	ATCAAGATAGTATTTTGTGTAGTTCAGACATGATCTATCTATTTTAAAGTATTT	535
Db	4227	ATTATATGTAGATTTAAAGAGTATTTATTTATTTATTAATAGTATTTAT-AGTATGTT	4285
Oy	536	TAGGTTGCTAATAGCTTATPAAAAATPAAAGAGGAAAAACATGATPAAAAAGTCTAT	595
Db	4286	TATATTTTATGTATATTATATTATTTATAGTAAGTATATTTTATATATGATTTATAT	4345
Oy	596	TAAAGTCAACCGTTTATAGTAAGAAATPACATTTTATCTCCAAA	640
Db	4346	TATGTATTAATATTTTGTGTTTAAATTTAAAGAAATTTTTTTTAA	4390

```

RESULT 12
US-10-312-841-2/c
Sequence 2, Application US/10312841
Publication No. US20030186277A1
GENERAL INFORMATION:
APPLICANT: Epigenomics AG
TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
FILE REFERENCE: E01/1208/WO
CURRENT APPLICATION NUMBER: US/10/312,841
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 2
LENGTH: 3673778.
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:
NAME/KEY: unsure
LOCATION: (379615)
US-10-312-841-2

Query Match
Best Local Similarity 6.1%; Score 47; DB 13; Length 3673778;
Matches 134; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
Pred. No. 1.1e+02;

QY 461 TAAGCAATAACCAATCAAGATAGTAGTTTGTAGTACAGACAGATACATCTATCCTA 520
DB 1227330 TATACCAATTAATAATACGAAAAAAATTCATTTAAAACTTACATTTAATACAAACAA 1227271
QY 521 TTTTATAGTATTTAGGCGTGCATAATAGCTATATAATAATAAGAGAGAAAAACAT 580
DB 1227270 AAACAAAAAAATTTATTTTAAACATAAAAAATTAATTAATTAATAAAAACTAA 1227211
QY 581 GATAAAAAAGTCATTAAAGCTCAACGTTTTTATAGTAAGATACAAATTTATCTCCAA 640
DB 1227210 AATAAAAAAATTACTTAACCAAAAAATTTAATACATTAACATTTTATACCACTAC 1227151
QY 641 CGATAAACGAGTTTACTGATATACTCAAGCATTGAGACTGAGCTAGATAAAATAAAGT 700
DB 1227150 ACTCCAACTAAACAAACAAACAAATACCTTTCACAAAAAATAAATAAATAA 1227091
QY 701 TTTTGGAAACAGTTACTACTAGCTAATCCATAACTCA 739

```

```

DB      1227090  TTAATAAAAAAAAAAACTACCTCTCTAAGTAACTTAACTTTA 1227052

RESULT 13
US-10-311-455-844/c
; Sequence 844, Application US/10311455
; Publication No. US20030143606A1
GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 844
; LENGTH: 5303
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-844

Query Match      6.1%; Score 46.8; DB 13; Length 5303;
Best Local Similarity 49.6%; Pred. No.10;
Matches 120; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY      520  ATTTTAAAGTATTTTAAAGGTTGCTAAATAGCTTATTAATAAATTAAGAGAAAAACA 579
DB      4961  ATTTTAAATAATTTCTCTTTTAAATTAATTAATTCGTAATAAATAATTAACCTCAACTCAC 4922

QY      580  TGATTAATAAGTTCATTTAAAGCTCAACCGTTTAAAGTAAGAAATACAAATTTTATCTCCAA 639
DB      4921  AATTAAACATATCTTTTAAATTTAACTTTTATATTTTAAATTAATTTTAACTTAAAT 4862

QY      640  ACGATAAACGAGTGTACTGTAATTAATCAAGTCAATGTAAGTCAAGTAAATTAAG 699
DB      4861  AAAATTAACACTCTCTCTTATTAATTAACATATTTCTTTTAAATTTTAAACAAAATTTAAT 4802

QY      700  TTTTGTGACAGTACTACTAGCTAATCCTAACCTCATATGATGTATGCAAAATATA 759
DB      4801  CTTTAAATTTACAAATTAATAAATAATTTTAAATTAATTAATTAATTTTCTCAATATA 4742

QY      760  AT 761
DB      4741  AT 4740

RESULT 14
US-10-311-455-959/c
; Sequence 959, Application US/10311455
; Publication No. US20030143606A1
GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30

```



```

? PRIOR APPLICATION NUMBER: DE_10043826.1
? PRIOR FILING DATE: 2000-09-01
? NUMBER OF SEQ ID NOS: 2424
? SEQ ID NO 959
? LENGTH: 17294
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURES:
? OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-959

```

Query Match	6.1%	Score 46.8	DB 13	Length 17294
Best Local Similarity	47.3%	Pred. No. 16		
Matches 141, Conservative	0	Mismatches 157	Indels 0	Gaps 0

QY		421	TGTAACATGAAGAAACAGCAACTGTGCATTGTGTGATTCACGTACGAAACAAATAACCAATACA	480
Db		8360	TTTTTAAATAAAACCAAATATATATCTTTCCAAATCTATATAAATTATATAAATCAAAAAATA	8307
QY		481	AGGATAGTATTTTGTTAGTCAGACATGAGATACATACCTATTTTATATAAGTATTTTAGGG	540
Db		8300	AATAAAAAAAAAAAAATTTTATATCCCTATTCCTAATTAATATATATTAATAAAAAATATATCTCTATA	8241
QY		541	TTGCTAAATAGCTATATAAAATPAAAGAGAGAAAAAACGTATAAAAGTTCATTTAAG	600
Db		8240	TATTTTAAATCTCTAAAATAACTATATAAATCCAAAAATAATATATAAAATTTTAAACAATA	8189
QY		601	CTCAACCGTTTTTGTAGTAAGAAATACAAATTTTATCTCAAAACGATTAACGGAGTTTACTG	660
Db		8180	CTCAAAACTTATCTCTAAATCAAAATATTTAATAAAAAACATTACTAATTAACCTATACAA	8123
QY :		661	AATATATCTCAGCATTTGAGACTGACGTGAAGTAAAAATTAAGTTTTTTTGGACACGTACT	718
Db		8120	AATCAATBAAAAACCTTCCAATTAATATATAAATCCACCTAACTCTTCAACATTAATCT	8063

```

RESULT 15
US-10-311-455-2189/c
; Sequence 2189, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2189
; LENGTH: 7403
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-2189

```

Query Match	6.1%	Score 46.6	DB 13	Length 7403
Best Local Similarity	48.9%	Pred. No. 13		
Matches 158; Conservative	0	Mismatches 159	Indels 6	Gaps 1

QY	4	TGAGGAATATTTAAATTTATTTAAATCTGATTTCTAGTCTCGAATATTAATAG	63
QY	4		
Db	4450	TCATTAAATTCCTGCATTTATTAACCTTTTACATTTATTTATCTTTTAAATACAC	4399
QY	64	ATAGCTTATTTAGCTCTTGAACATCTGAATGACCTAGCTTTATTAACATATACACAT	123

Accession	Sequence	Position
Db	4390 ATTTTATTTTATCTTTTAAAAATACATATATTTTATCTTTTAAAAATACGATCAT	4331
Oy	124 AGAAACCTTAAACAATCTAAAAAGTC-----TTAATTCATCTTGAAAGATTCGT	177
Db	4330 ATATTACTTTAAATATATACATCTTCATCTAATTATTTCTTTAAATTTTAACTATATATTA	4271
Oy	178 AATATATTTATTTGCGATTAACGAGCATATTAACGGCTCTGATTAATTCGAAGTT	237
Db	4270 AATATATTTTATCTCTCCAAATTTACAATATAAAAACAAAACTTAAAAAATCTTAAAT	4211
Oy	238 GTTGAATACATGATGTTTCGTTCGAAGGACATACAAAATTAATTTAAAGAGGACCTCAA	297
Db	4210 AAAAAATATAAAAAAATATATTCATATACAAAAAATATTAATAATTTTAAACTTACTTTA	4151
Oy	298 ATGAGTACAAAAGATTTTAACCTT 320	
Db	4150 AATATATATATCAATCAATCACTT 4128	

Search completed: January 12, 2004, 14:02:55
Job time : 928 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2004, 11:20:01 ; Search time 2084 Seconds
(without alignments)
8945.074 Million cell updates/sec

Title: US-10-082-618-8

Perfect score: 767
Sequence: 1 agtcgacgaatcattatcaataa.....tcgcagaataatcgtcgt 767

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlum:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82.4	10.7	1101	29	CNS00EVL
2	74.4	9.7	1101	29	AL069706 Drosophila
3	71.6	9.3	781	29	AL063921 Drosophila
4	70.6	9.2	1200	13	BX145762 Dario rer
					BX437758 BX437758

Result No.	Score	Query Match	Length	ID	Description
5	67.2	8.8	1200	13	BX415878
6	67.2	8.8	1201	9	AL536104
7	64.8	8.4	1056	13	BX415058
8	64.6	8.4	1201	9	AL536104
9	63	8.2	1200	13	BX415878
10	62.6	8.2	1201	29	CNS00390
11	61.8	8.1	1201	9	AL565455
12	61	8.0	1200	13	BX437758
13	59.4	7.7	1201	9	AL543368
14	59.2	7.7	928	29	CNS00DKY
15	59.2	7.7	960	13	BX346155
16	59.2	7.7	1201	13	BX355654
17	59	7.7	987	29	CNS014PQ
18	59	7.7	1098	13	BX377526
19	58.8	7.7	1201	13	BX361152
20	58.4	7.6	804	28	B12681
21	58.2	7.6	928	29	CNS00DKY
22	58	7.6	1061	29	CNS015LM
23	57.8	7.5	985	13	BX436838
24	57.6	7.5	1200	29	CNS016CO
25	57.4	7.5	893	29	CNS013XE
26	57.4	7.5	1101	29	CNS00215
27	57.4	7.5	1101	29	CNS00FYG
28	57.2	7.5	854	9	AL522840
29	57.2	7.5	1064	13	BX361825
30	57.2	7.5	1225	29	CC238324
31	56.8	7.4	1169	29	CNS006KHQ
32	56.8	7.4	1201	9	AL532464
33	56.4	7.4	1201	29	CNS0167M
34	56.2	7.3	1056	13	BX415058
35	56	7.3	1101	29	CNS0182P
36	56	7.3	1225	29	CNS0161D
37	55.8	7.3	873	29	CNS024MS
38	55.8	7.3	994	13	BX414650
39	55.8	7.3	1201	13	BX420717
40	55.6	7.2	986	13	BX366417
41	55.6	7.2	1124	13	BX436282
42	55.2	7.2	1167	29	CNS07360
43	55	7.2	996	29	CNS00FUR
44	55	7.2	1101	29	CNS00BO1
45	54.8	7.1	1201	13	BX422260

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

CNS00EVL. 1101 bp DNA linear GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29623 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL069706
AL069706.1 GI:4949849
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Bukariyola; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Oseegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of

QY 470 AACCAATCAAGAGTAGTATTTGTTAGTCAGACATGATCTATCTATTTTAA 529
DB 683 RRDRTWRTDAMWMDAAWMTTDDTDMDRDRDRRKARRRRTTAAADWMTWKA 624
QY 530 GTTATTTAGGTTGCTTAATAGCTATATAAAATTAAGAGAGAAAAACATGATTAAG 589
DB 623 DWAKDWTDRDRMDRMAADWTDRKADRWAKRARMARDRRAARADRRWTKGT 564
QY 590 TTTATTTAAGCTCAACGGTTTGTAGTAAGAAATCAATTTATCTCCAAAGATTAAG 649
DB 563 TTTATTTAAARAAMWAMWAMATTTATTTTWTTTTWTTTTWTAAWMAA-- 506
QY 650 GAGTTTACTGATATCTCAAGTCATGAGCTGATGAATAAAATTAAGTTTGTGA 709
DB 505 -WTATTAATTAATAAAWAAAAWATTTTWTAAWMTAAWMTWTTTWT 447
QY 710 ACAGTTACTACTAGCTATCTCTAAACTCTATGATGTTATGCAAAATTAATGC 763
DB 446 WAAATTTTWTWTTWMAATTAATTTTWTWTTWTTWTTWTTATATKCC 393

RESULT 3
BX145762/c BX145762 781 bp DNA linear GSS 28-JAN-2003
LOCUS Danio rerio genomic clone DKEY-108G13, genomic survey sequence.
DEFINITION BX145762
ACCESSION BX145762
VERSION BX145762.1 GI:29777115
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
REFERENCE Humphray, S.J., Huckle, E. and Durham, J.L.
AUTHORS Direct Submission
TITLE Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
JOURNAL Campy, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphrey@sanger.ac.uk unpublished
COMMENT This sequence was generated from the SP6 end of BAC 108G13. 108G13
is part of the Daniokey BAC Library created by R. Plasterk and N.V.
Keygene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES
source location/Qualifiers
1..781
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-108G13"
/issue_type="Testis"
/note="vector pindigobAC-536"

BASE COUNT 256 a 55 c 45 g 424 t 1 others

ORIGIN
Query Match 9.3%; Score 71.6; DB 29; Length 781;
Best Local Similarity 46.9%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 224; Conservative 0; Mismatches 254;

QY 9 AATATTTATATTTTATTAATCTGATTTTCTGATTCCTGATATATAGATAG 68
DB 710 AATATTTATATGATTAATACATTAACAGTATTAATAATTAATTAATTAATAC 651
QY 69 TTTATGAGCTTAGACATCTGATGATGACCTGCTTAATACTATGACATGA 128
DB 650 AATTAATTTTATTAATAAATAAGTAATAATTAATTAATTAATTAATTAAT 591
QY 129 CATTAAACAATCTAAACAGCTTAATCTATCTGAGAAAGTATGTAATATAT 186
DB 590 TATTAATATATATTAATTAACGTTAAACAGTATTAATAATTAATTAATTAAT 531
QY 189 TGTGATTAACGAGCATATAACGCTGATTAATTAATCTGAGTTGTTAGATCA 248
DB 530 TAAATTTAATAAATAAAGTTAATAATTAATTAATTAATTAATTAATTAATTA 471

QY 249 TGATTTGCTCGAAGAACTACAAAATTAATTAAGAGGCACTCAAAATGATGACAA 308
DB 470 TATTAAGCTTAACAGTATTAATAATTAATTAATTAATTAATTAATTAATTA 411
QY 309 AGATTTTAACTGATTTGTTGATCTGTTTGAAGAAAGATTCAGTGATCAACAGCAT 368
DB 410 TAAATAATTAAGTTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 351
QY 369 TACAGTATTTGCTGATGATTAACCCGGTTGTAACAGAGCTGATGAGTTGTAACAT 428
DB 350 TAAAGTATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 291
QY 429 GAAAACAGCACTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 486
DB 290 AGTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 233

RESULT 4
BX437758 BX437758 1200 bp mRNA linear EST 15-MAY-2003
LOCUS BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP0081B01
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX437758
VERSION BX437758.1 GI:30773605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0CAP008CA01QPL.

FEATURES
source location/Qualifiers
1..1200
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP0081B01"
/issue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT 515 a 30 c 71 g 310 t 274 others

ORIGIN
Query Match 9.2%; Score 70.6; DB 13; Length 1200;
Best Local Similarity 30.9%; Pred. No. 0.16; Indels 5; Gaps 3;
Matches 232; Conservative 157; Mismatches 357;

QY 11 TATTTAATTAATTTATTAATATCTGATTTTCTAGTCTGAATPAATAGATAGGTT 70
DB 372 TATWMMWDDDDAKWADAGATWTTTWTWBRGARGARKDGTDDPAADWMTWAAW 431
QY 71 TATGAGCTTAGACACTGATGATGACCTGCTTAATACTATACAGACATGAAGA 130
DB 432 AAWMDKRTTTTTTTTKRRRAAAATTTTTTTTTTTWMAAAAARAAATWTTWRAA 491
QY 131 TTAACAATCTAAACAGCTTAATCTATCTGAGAAAGTATGTAATTAATTTG 190
DB 492 AAAAAMWTTWMAAGRARGATTTTTTTTTTAAWMMWGAGRTAAMWTTWTTTTTTT 551

Oy		191	TGCAATACGCCA-GCATATAAACGGCTCGTGTAAATTTCGAAGTTTGTAACAT	249
Db		552	TAMWATWTATATTTTTTTWWTAAAAAASAGAKATTTTICKTTTATTWKRGAGATWTT	611
Oy		250	GATTTTCGTCGAGGAACTACAATAATTATTAAGAGGCACTCAAAATGATCAAAA	309
Db		612	TTTTTTTWWARGMKWMTWTTTTTTTTTTTTTTTTTTTTTTTGTGTTTWWBDGARRTAKATK	671
Oy		310	GATTTTAACTGGATTGCGTATCCTGTTCCGAAGAAGATTACGGTCATCCACGCATT	369
Db		672	WTTTWTWTTAAWAGGAARKRATWTTWAAAAWAAGAAAAARGAAGRAAWTTTTTTT	731
Oy		370	ACAAGTATTTGCTATGTACACCCGCTGTAAACAGAGCTCTGATGGTGTAAACATG	429
Db		732	ATKGARGATWTTTAAWMTATPRARGAGADTTTTTATNMTATTTTTTTWWMAAGATD	791
Oy		430	AAAAACGCAACTGTCTATTGTATGATTCACGTAAAGAAATACCAAATCAAGATAGTA	489
Db		792	KAAAAAAMWMTTWTAAAAAAATTTWWAGAPAAK-ATWTAWAAAAAATAAAARAWA	850
Oy		490	TTTTGTAGTCCAGCATGATGATCATCTATCTTTTATTAAGTATTTAGGGTGTAAAT	549
Db		851	ATATATWTTTATPATATKAKAPAAAAATATAAARARRWKGAATAAAAAATTAATAW	910
Oy		550	AGCTTATTAATAATAAGAGAGAAAAACATGATTAATAAGTTCATTAAAGCTCAACGT	609
Db		911	WTTWTATAAAAAAATTTWAAWMBABBAWMAWMAAAGABAAGAMAA---WATWTT	967
Oy		610	TTTTTAGTAAATATCAATTTTATCTCCAAACGATAAACGAGTTTATCTGAATACTC	669
Db		968	TTTTTAATATRWMAAAAAAAMTTTTTTTTTWWMAAAAAAMWDTTATATWTAWTATAGAW	1027
Oy		670	AAGCATGTGACACGTAACTAAATAAGTTTTTTTGAAACAGTACTACAGTAAATC	729
Db		1028	AAATARDTWPAAATTAATYADARRAAGAAWAIKDGAGATTAATTTTTTTTTTTTWT	1087
Oy		730	CTAAACTTATGATGTTTATGCAAAATATAA	760
Db		1088	AAAAAAAAAAAAATTTTAAAAAANAANA 1118	
RESULT 5				
BX415878/c				
LOCUS			1200 bp mRNA linear EST 15-MAY-2003	
DEFINITION			BX415878 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008Y104	
VERSION			5-PRIME, mRNA sequence.	
KEYWORDS			BX415878 BX415878.1 GI:30765550	
SOURCE			EST.	
ORGANISM			Homo sapiens (human)	
REFERENCE			Homo sapiens	
AUTHORS			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
JOURNAL			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
COMMENT			1 (bases 1 to 1200)	
			L.I.W.B., Gruber,C., Jeesee,J. and Polayes,D.	
			Full-length cDNA libraries and normalization	
			Unpublished	
			Contact: Genoscope	
			Genoscope - Centre National de Sequencage	
			BP 191 91006 EVRY cedex - France	
			Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr	
			Laboratory was constructed by Life Technologies, a division of	
			Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :	
			http://fulllength.invitrogen.com/ Invitrogen Corporation 1600	
			Faraday Avenue Genoscope sequence ID : CS0CAP008BER02DP1.	
FEATURES				
Source			1..1200	

[illegible]

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1201)
REFERENCE	Ll, W. B., Gruber, C., Jesse, J., and Polayes, D. Full-length cDNA libraries and normalization unpublished
AUTHORS	
TITLE	
JOURNAL	
COMMENT	On Feb 13, 2001 this sequence replaced gi:12799597.

BASE COUNT	359 a	119 c	146 g	409 t	168 others
ORIGIN					

Query Match	8.8%;	Score 67.2;	DB 9;	Length 1201;
Best Local Similarity	33.8%;	Pred. No. 0.52;		
Matches 185;	Conservative 94;	Mismatches 266;	Indels 3;	Gaps 1

QY	119	ACATATGAGAAACATTAACAAATCTAAACAGCTTAAATCTAATCTGAGAAAGATATGTGGA	178
Db	651	AMTATATTDMMWTTATATAADDBTATATAAATTTTATATWTTTATATTTTWTWTAATTTT	710
QY	179	ATAATATATTATGTCGATPACGGAGCATATATAACGGCTCTGATTAAATTCGAAATTG	238
Db	711	ATATAAAATATATATATATWATTTTWTWTTWTTATATACAMMTWTTWATATATTTTAAATTTT	770
QY	239	TTAGATACATGATTTGTTGTCGAGAAACTACAAAATAATATATAAGAGGCACTCAGAAA	298
Db	771	ATAATTTATRWATATATWTTTWTWTTTWTWATTAATWATATWTTAAWTTTWTATTT	830
QY	299	TGAGTACAAAGAATTTTAACTGGATTTGGTATCTGTTTCGAGAAAGAATTCAGGTCAT	358
Db	831	WTATWATAAAMWTTTWTATATWTTATATTTTWTWATAAAATTTATTT--WTTTATATTTAAW	887
QY	359	CACCAACGATTACAGATATTTGCGATGTGACACCCGGTTGTAAACAGAGAGCTCGATGG	418
Db	888	TTTATWTTTATTTATTAATAAMTTTWTATAMAMWATTTTWTWMAATTTAGTATAMAMAMATA	947
QY	419	GTTGTACATGAAAAACAGCAACTGTCTATTGTGATTCAGTAAACAAATACCAAAATC	478
Db	948	TATATATAMACMRTAAATAMWMTTAAWAAWTATATWTTAAAAAAMAWAATWMAAATATWTTT	1007
QY	479	AAAGATATGTAATTTTGTAGTTACACATCGATACTATCCATTTTATATACTATATTAG	538
Db	1008	ATMATTTWTTTWTWTTWTTWATWSTATWATATAAAAAATAMTWTTTTTTAAATWATAMWATW	1067
QY	539	GTTGCTAAATAGCTTATPATAAATAAAGAGAGAAAAACATGATPAAAAAGTTCATTTTA	598
Db	1068	TAAAMAMAMRTWTTTATABAAAAMAMWATTTTWRTRWMAAAAAAABAAAAATADMAATDW	1127
QY	599	AGCTCAACCGTTTATGTAGTAAGAAATACAATTTATCTCCAAACGATPAAACGAGTTTAAAC	658

Db 1128 WMTTGGATATATTATTAATWTTTADAMDATTTAAATATTAATTAATTAATTAATA 1187

QY 659 TGAATATA 666
| | | | |
Db 1188 TAAAKATA 1195

RESULT 7	LOCUS	DEFINITION
BX415058	BX415058	1056 bp mRNA linear EST 15-MAY-2003
	BX415058	Homo sapiens THYMUS Homo sapiens CDNA clone CS0CAP004YG19 3'-PRIME, mRNA sequence.

FEATURES

Query Match	8.4%	Score 64.8	DB 13,	Length 1056,
Best Local Similarity	34.7%	Pred. No. 1.2,		
Matches 210, Conservative	87,	Mismatches	299,	Indels 9,
				Gaps 1,

```

QY      52  AATAATATAGAGATAGTTTATTGAGCTTACAGATCACTGGAAATGACCTACTCTTAATAC 111
Db      322  AACCAACAATAGTTATATAAACTTGGCTGCTGTAATTTTATATATATATATTTTATA 381
QY      112  TATACGACATAGAAACATTAACTAACTCTAAATCTTATCTTGAGAAAGT 171
Db      382  TATATATATMAAAAAAAAAAAAAAAAAAAAMTKATATATAMTTTKTTTTTWTATTT 441
QY      172  ATTGGTATATATATTATTGTCGATACGCGAGCATATAAACGGCTCTGATTAATCTG 231
Db      442  ATAAAAAARWATATATTTAAAAAAAAMTTTWWAAAAAAAATTTTATWMAAATTTWA 501
QY      232  AAGT-----TTGTGATCAATGATTTTGGTGAAGAACTCAAAATTAATAT 282
Db      502  AAAAAAAMATTWDDKKKKMTATATAAAAAAAAAAAAAAAMMAATATATATATATTTAT 561
QY      283  AAGGAGGCACTCAAAATGAGTACAAAAGATTTTAACTGTGATTTGGTATCTGTTCCAG 342
Db      562  AKATATATAAAAAAAAAAAWATATATATAAAAAAAAAAAAAAAMTTTGGWAAWATGTTTTT 621
QY      343  AAAGATTCAAGTGCATCACACGCAATTACAAATTTTGGCTATGTACACCGGTTGTAA 402

```

```

Db      622  TTWTTTWTWMAATATTTTATADATPTKTTTTTWTATDMDATWAAWTATWTWRTGKA 681
QY      403  ACAGAGGCTCGATGGCTGTGAACGTAAAACAGCACTTGTCATTGATGATTTACGTA 462
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      682  GWMMAAAAKTKATRAKTKTKTAMMAAAAAAAAAAAAAAAAAAAAAAAAAAAWMT 741
QY      463  AGCAATPACCAAAATCAAGAATAGTATTTTGTTAGTTCAGACATGATACTACTCATTT 522
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      742  TTTTWTWTWMAAAAAAAAAATNNNNNNNTTAAAMWMTTTTTTTTTTTTTTMMWDDDDDD 801
QY      523  TTTATPAGTATTTAGGGTGTCTAATATAGCTTATATATAAGAGAGAAAAAATCA 582
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      802  DTTTTTTTTTTTTTTTTTMMWMTMMWMTMMWMTMMWMTMMWMTMMWMTMAAAAAAAT 861
QY      583  TAAAAAGTTCATTAAAGCTCAACCGTTTTTGTAGTAAGAAATACATTTTATCTCCACG 642
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      862  TTTWTAAATTTTATAMWMAAAATGTTTRWAMAAAAAAAAAAAAAAAAAATTAAAMAAWMTT 921
QY      643  ATAAA 647
        : : : : :
Db      922  WMAAA 926

```

RESULT 8	AL536104/c	LOCUS	DEFINITION
AL536104	1201 bp	mRNA	linear
AL536104	Homo sapiens FETAL BRAIN	Homo sapiens	CDNA clone
CSDDP022YC18	5-PRIME, mRNA sequence.		

ACCESSION	AL536104	GI:31260974
VERSION	AL536104.2	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
TITLE
AUTHORS
JOURNAL
COMMENT

1 (bases 1 to 1201)
L4, W.B., Gruber, C., Jeese, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:1279597.

Contact: Genoscope
Genoscope - Centre National de Séquençage
BP 191 91006 Evry cedex - France
Email: seget@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF022BB09QPt.

```

FEATURES
SOURCE
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF022XC18"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NciI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoR sites of the pCMVSPORT 6
vector. Library was not normalized."
359 a 119 c 146 g 409 t 168 others
BASE COUNT
ORIGIN

```

[illegible]

QY	196	AAGCGAGCAATAA	AAACGGCTCGATTAATCTGAAAGTTGGTAGATCAATGATTC	255							
Db	1128	WMTATWTHATATTT	TTTTTTTTTTTWTWTA	WMAATATWMTWTTTTTTAA	MAWA	YMTWMTTT	1069				
QY	256	GTTGGAAGAACTA	CAAAATAAATTTAT	TAGAGAGCACTCAAAATGAGTACAAAGATTTT	315						
Db	1068	AAAWMTATWTA	MTTAAAAA	MAATWTTT	TTTATWATAS	NAWMAA	MAWA	MAATWTA	1009		
QY	316	AACATGGATTTG	TAATCTGTTTGCA	AAAAAGTTCAAGTGCATCAACCA	CGCATTAAGT	375					
Db	1008	TAAAWATATTT	WMTWMTWTTTTTTT	TA	MAATWMTWMAA	MTWTA	MTTAA	YMCWTA	MTAT	949	
QY	376	ATTTCGCTATG	CACCCGGTTG	TAAACAGAGCTCGATGGGTTTAA	CATGA	AAACA	435				
Db	948	ATTW---	TWMTWMTAKCTA	TATWATTTWMTWMAA	ATWMTWTA	MAAA	MTTATA	ATAA	TA	MAA	892
QY	436	GCAACTGTCTA	CTGTAGTATTC	ACGTAGCAATAACCA	AATCAAGATATG	TTTGT	495				
Db	891	TAAATWTTAA	ATTTAAAA	MAATTAATTTTA	MAAAAAA	TATWMAA	ATWMAA	MTTWTW	MTA	832	
QY	496	TAGTTGACAT	GATCTACTATCCATTTT	TATATAGTATTTAGGCTGCTAAATAGCTA	555						
Db	831	WATATWAAAA	MTTATWMTATTTAT	TATWMAA	MAWA	MAA	MTATWMTAT	TA	MAA	MTA	772
QY	556	TAAAAATTA	AGAGAAAAA	ACATGATATAAAAGTTCA	TATAAGCTCAACCGTTTTAG	615					
Db	771	TAAAAATTT	AAAAATATATW	AAWMTATATATWMAA	MAA	MAA	MTATATATATTTWTA	712			
QY	616	TAAAGAAAT	TACAATTTA	632							
Db	711	TAAATATW	MAAAAAATA	695							

RESULT 9	1200 bp	EST 15-MAY-2003
EX415878	mRNA	linear
LOCUS		
DEFINITION	THYMUS Homo sapiens	CDNA clone CS0C4P008Y104
	5'-PRIME, mRNA sequence.	

ACCESSION	BX415878
VERSION	BX415878.1
KEYWORDS	GI:30765550
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1200)	Li, W. B., Gruber, C., Jesse, J., and Polares, D.	Full-length cDNA libraries and normalization	unpublished	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Radway Avenue, genoscope_sequence ID : CSCAP008BE02QPL.

```

FEATURES
    source
        location/qualities
            1..1200
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0CAP008Y104"
                /tissue_type="THYMUS"
                /clone_id="Homo sapiens THYMUS"
                /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                    with a NotI-oligo(dT) primer. Five prime end enriched,
                    double-strand cDNA was digested with Not I and cloned into
                    the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                    Library was not normalized."
BASE COUNT
    428 a      70 c      85 g      447 t      170 others
ORIGIN

```


VERSION AL565455.2 GI:30549492
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jesssee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 16, 2001 this sequence version replaced gi:12916848.
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9232.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSODP005BH09NP1&cluster=9232.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSODP005BH09NP1.
 FEATURES
 source
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODP005Y018"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."
 BASE COUNT 550 a 82 c 71 g 313 t 185 others
 ORIGIN
 Query Match 8.1%; Score 61.8; DB 9; Length 1201;
 Best Local Similarity 30.2%; Pred. No. 3.5;
 Matches 219; Conservative 122; Mismatches 384; Indels 0; Gaps 0;
 10 ATATTGTAATTTTATTAATCTGATTTCTAGTCCGTAATATATAGATAGT 69
 448 AAT 507
 70 TTATTGAGCTTGAACATCTGATGACCTGCTTATATCTATGACATAGAAC 129
 508 AAT 567
 130 ATTAACAATCTAAACAGCTTAACTTATCTTGAGAAAGATTGGTATATATT 189
 568 AAT 627
 190 GTCGATTAACGAGCATATATAACGCTGATTAATCTGAAGTTGTTGATACAT 249
 628 AAT 687
 250 GATTTCGTCGAAGAACTACAAATTAATTAAGAGGAGCTCAATGATGACAAA 309
 668 AAT 747
 310 GATTTAACTGATTTGATCTGTTTGAAGAAAGATTGAGTGATCACCAGCAT 369
 748 AAT 807
 370 ACAATATTTGCTATGTCACCCGCTTAAACGAGCTGATGGTGTATACATG 429
 808 AAT 867
 430 AAAACAGCACTTGATCTGATGATTCACGTAGCAATTAACAATCAAGATATGA 489

DB 868 AAT 927
 QY 490 TTTTCTTACTGACGATGATGATCTATCTTATTTTATAGTATTTGCTAAT 549
 DB 928 AAT 987
 QY 550 AGCTTAT 609
 DB 988 AAT 1047
 QY 610 TTTTGAAGAAAT 669
 DB 1048 AAT 1107
 QY 670 AAGTATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
 DB 1108 AAT 1167
 QY 730 CTAA 734
 DB 1168 AATAT 1172
 RESULT 12
 BX437758/c 1200 bp mRNA linear EST 15-MAY-2003
 LOCUS BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008Y01
 DEFINITION 5-PRIME, mRNA sequence.
 ACCESSION BX437758
 VERSION BX437758.1 GI:30773605
 KEYWORDS EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1200)
 Li, W.B., Gruber, C., Jesssee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSOCAP008CA010P1.
 FEATURES
 source
 1..1200
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOCAP008Y01"
 /tissue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."
 BASE COUNT 515 a 30 c 71 g 310 t 274 others
 ORIGIN
 Query Match 8.0%; Score 61; DB 13; Length 1200;
 Best Local Similarity 32.0%; Pred. No. 4.6;
 Matches 233; Conservative 127; Mismatches 360; Indels 9; Gaps 2;
 QY 3 TTGAGAT 62
 DB 1197 TTAT 1138
 QY 63 GATAGTTTATGAGCTTATGACATCTGATGATGATGATGATGATGATGATGAT 122

```

Db 1137 TTTT... 1078
Qy 123 TAGA... 182
Db 1077 AAAA... 1018
Qy 183 TATTA... 242
Db 1017 WATA... 958
Qy 243 ATACA... 302
Db 957 TCTYT... 898
Qy 303 TACAA... 362
Db 897 TTTCA... 842
Qy 363 AGCA... 422
Db 841 TTTT... 782
Qy 423 TAACA... 482
Db 781 AAAA... 722
Qy 483 GATAG... 542
Db 721 YCTTC... 662
Qy 543 GCTAA... 602
Db 661 CHYMA... 607
Qy 603 CAACG... 662
Db 606 CCTVM... 547
Qy 663 TATAC... 722
Db 546 AAAA... 487
Qy 723 GCTAT... 731
Db 486 AATAT... 478

RESULT 13
AL543368 1201 bp mRNA linear EST 31-MAY-2003
LOCUS AL543368 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CSOD1001YF24 5-PRIME, mRNA sequence.
ACCESSION AL543368
VERSION AL543368.2 GI:31265215
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li W.B., Gruber C., Jessee J. and Polayer D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12875846.
Contact: Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9562.r For
more information about this cluster, see
http://www.genoscope.cns.fr/

```

```

FEATURES
source:
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="CSOD1001YF24"
  /tissue_type="PLACENTA COT 25-NORMALIZED"
  /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
  /note="1st strand cDNA was primed with a NotI-cligo(dT)
primer. Five prime end and cloned, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMSPORT 6 vector. Library was normalized."
BASE COUNT
453 a 178 c 231 g 174 t 165 others

Query Match
Best Local Similarity 34.7%; Pred. No. 8;
Matches 124; Conservative 66; Mismatches 167; Indels 0; Gaps 0;

Qy 232 AAGTT... 291
Db 807 AAADG... 866
Qy 292 CTCAA... 351
Db 867 ATWAA... 926
Qy 352 GGTGA... 411
Db 927 GKTGT... 986
Qy 412 CTGAT... 471
Db 987 DKKT... 1046
Qy 472 CCAAT... 531
Db 1047 AAAA... 1106
Qy 532 TATTA... 588
Db 1107 AAAA... 1163

RESULT 14
CNS00DKY/c 928 bp DNA linear GSS 04-JUN-1999
LOCUS CNS00DKY Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BACR27A24 of RPCT-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL071865
VERSION AL071865.1 GI:4948170
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 928)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: segref@genoscope.cns.fr)
Web: www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila

```

melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source

Location/Qualifiers

1..928
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR27A24"
/clone_lib="RPCI-98"
/note="end: T7"

BASE COUNT 262 a 70 c 84 g 321 t 191 others

ORIGIN

Query Match 7.7%; Score 59.2; DB 29; Length 928;

Best Local Similarity 31.7%; Pred. No. 9.2; Mismatches 135; Indels 0; Gaps 0;

Matches 97; Conservative 74; Mismatches 135; Indels 0; Gaps 0;

26 TTAATATCTGATTTCTAGTCCGTAATATAGAGAGTTTATGAGCTTAGAC 85

825 TMMNNHHNNHHNNHHNNHHNNHHNNHHNNHHNNHHNNHHNNHHNNHHNN 766

86 ATAAGTGAATGACCTGATCTTAATACATGACAAATGAAACATTAACTTAA 145

765 WMMNN 706

146 CAGCTTAATCTGATCTGAGAAAGATGTAATATATTTGCGATAACGAGCA 205

705 AATTTTWTATTAATTTTAAATTTTATTTTATTTTATTTTATTTTATTA 646

206 TAATTAACGCTCTGATTAATTTGTAAGTTTGTATACATGATTTGTTGCA 265

645 AAMAMAMAMATTTATATATATATTTTATTTATTTATTTATTTATTTAT 586

266 ACTACAAATTAATTAAGAGCACTCAAAATGATCAAAAGATTTTAACTTG 325

585 ATAMATTTTAAATTAAGAGCACTCAAAATGATCAAAAGATTTTAACTTG 526

326 TGGTAT 331

525 TTTTWW 520

RESULT 15

BX346155/c 960 bp mRNA linear EST 05-MAY-2003

LOCUS BX346155 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED

DEFINITION Homo sapiens cDNA clone CS0DJ010Y14 5-PRIME, mRNA sequence.

ACCESSION BX346155

VERSION BX346155.1 GI:30373068

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 960)

14.W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr

Library was constructed by life technologies, a division of

Invitrogen. Contact: Peng Liang Email: filang@lifetech.com URL:

<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0AJ010BB07QPI.

FEATURES

Source

Location/Qualifiers

1..960
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ010Y14"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

BASE COUNT 284 a 100 c 80 g 366 t 130 others

ORIGIN

Query Match 7.7%; Score 59.2; DB 13; Length 960;

Best Local Similarity 35.1%; Pred. No. 9.1; Mismatches 376; Indels 10; Gaps 2;

Matches 259; Conservative 92; Mismatches 376; Indels 10; Gaps 2;

9 AATATTAAATTTTATTAATATCTGATTTCTAGTCCGTAATATATGAGATGG 68

940 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 881

69 TTTATGAGCTTACACATCTGTAAGTACCTGATTAATTAATTAATTAATTA 128

880 AATTTTWTATTAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTA 821

129 CATTAACAATCTAAACAGCTTAATTTCTATCTGAGAAATGATTTGTAATTA 188

820 WTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 761

139 TGTGATTAACGAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 248

760 TTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 701

249 TGATTTGCTGGAAGCACTCAAAATTAATTAATTAATTAATTAATTAATTA 308

700 ATTWTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 649

309 AGATTTTAAGTGTGATTTGTAATCTGTTGCAAGAAAGATTCAAGTGCATCA 368

648 ATTWTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 589

369 TACAAGTATTTGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 428

588 MNN 529

429 GAAAAACAGCACTGCTATGATTAATTAATTAATTAATTAATTAATTAATTA 488

528 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 469

489 ATTTTGTAGTTCAGACATGATCTATCTATTTTATTAATTAATTAATTAAT 548

468 GAACATTTGCTGCTGGAATTAATTAATTAATTAATTAATTAATTAATTAAT 409

549 TAGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 606

408 TAAAMMMTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 349

607 CGTTTGTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 666

348 GGTGAATATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 289

667 CTCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 726

288 MAAAMATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 229

727 ATCTTAATCTATGAT 743

228 GTCAATGCTCTTCAT 212

Search completed: January 12, 2004, 12:04:59
Job time : 2089 secs

THIS PAGE BLANK (USPTO)